Hidden-Markov Methods for the Analysis of Single-Molecule Actomyosin Displacement Data: The Variance-Hidden-Markov Method

David A. Smith,* Walter Steffen,[†] Robert M. Simmons,[†] and John Sleep[†]
*The Randall Centre for Molecular Mechanisms of Cell Function and [†]MRC Muscle and Cell Motility Unit, King's College London, Guy's Campus, London SE1 1UL, United Kingdom

ABSTRACT   In single-molecule experiments on the interaction between myosin and actin, mechanical events are embedded in Brownian noise. Methods of detecting events have progressed from simple manual detection of shifts in the position record to threshold-based selection of intermittent periods of reduction in noise. However, none of these methods provides a "best fit" to the data. We have developed a Hidden-Markov algorithm that assumes a simple kinetic model for the actin-myosin interaction and provides automatic, threshold-free, maximum-likelihood detection of events. The method is developed for the case of a weakly trapped actin-bead dumbbell interacting with a stationary myosin molecule (Finer, J. T., R. M. Simmons, and J. A. Spudich. 1994. *Nature.* 368:113–119). The algorithm operates on the variance of bead position signals in a running window, and is tested using Monte Carlo simulations to formulate ways of determining the optimum window width. The working stroke is derived and corrected for actin-bead link compliance. With experimental data, we find that modulation of myosin binding by the helical structure of the actin filament complicates the determination of the working stroke; however, under conditions that produce a Gaussian distribution of bound levels (cf. Molloy, J. E., J. E. Burns, J. Kendrick-Jones, R. T. Tregear, and D. C. S. White. 1995. *Nature.* 378:209–212), four experiments gave working strokes in the range 5.4–6.3 nm for rabbit skeletal muscle myosin S1.

INTRODUCTION

The actomyosin motility assay (Kron and Spudich, 1986) has spawned a number of methods of monitoring interactions between actin and myosin, using a microneedle (Kishino and Yanagida, 1988), a single optically trapped bead (Miyata et al., 1994), two optically trapped beads (Finer et al., 1994), and a hybrid bead-microneedle system (Ruff et al., 2001). Although experimental techniques have become increasingly sophisticated, signal-to-noise ratios are usually low and reliable detection of interactions ("events") can be problematic (Block and Svoboda, 1995; Mehta and Spudich, 1998). To measure the maximum working stroke of myosin, the force opposing movement should be minimized by using optical traps or microneedles that are much less stiff than myosin itself: the stiffness of myosin is ~1 pN/nm, and that of the measuring device is typically only 0.04 pN/nm. Consequently, the relative Brownian motion of myosin and actin before binding is large, typically ±10 nm rms, whereas the working stroke may be as low as 5 nm.

In the first direct determination of the working stroke, Finer et al. (1994) recorded the position of one of a pair of optically trapped beads while an actin filament stretched between the beads interacted with a myosin molecule bound to a surface (Fig. 1). Using filtered position data, they scored any clear stepwise transient displacement from the

baseline as an event and obtained a mean value of ~11 nm, which they took to be the value of the working stroke. However, Molloy et al. (1995) showed that this method excluded events that are hidden in the residual baseline noise; the weakly trapped actin-bead "dumbbell" is translated by Brownian motion by up to ±30 nm past the myosin molecule, which is presumed to be able to attach to any actin monomer presented to it. The observed bound displacement level includes both the displacement of the actin site before binding and the subsequent working stroke: the working stroke cannot be determined from an individual event as it is not possible to distinguish the pre-stroke displacement level from the noise. Thus the distribution of bound levels is governed by the same Gaussian function describing bead displacements in the absence of attachment, but with every point shifted by the working stroke. If the working stroke is constant and independent of displacement, its value can be defined as the difference in the means of these two Gaussian distributions. Molloy et al. showed that attachment events could be detected by changes in a variance record derived from the bead displacement record: the system becomes stiffer when myosin binds and the noise level decreases. On this basis the value of the working stroke was found to be about half the value measured by Finer et al. (1994), which was based solely on bound levels from the uppermost tail of the distribution.

Later work used threshold methods (Colquhoun, 1998) to detect single edges in a small window of the derived variance record from one bead (Veigel et al., 1998; Ishijima et al., 1998), or the covariance record from both beads (Mehta et al., 1997). The "mean-variance" method (Patlak, 1993) has also been used to analyze displacement records (Guilford et al., 1997; Tyska et al., 1999); this method does not
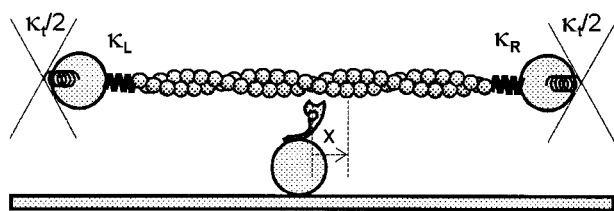
FIGURE 1 Schematic diagram of the actin-bead assembly moving in optical traps and in interaction with a single tethered myosin-S1 molecule. The mechanical properties of this system are represented by an inextensible actin filament and lumped elastic constants k for myosin, $\kappa_t/2$ for each trap, and κ_L, κ_R for the links between filament and each bead. When myosin binds to a monomeric actin site at distance x to the right, it is assumed to exert a left-directed force $k(x + h)$; if binding is most rapid when $x = 0$. For an unloaded filament, this force would produce a left-directed displacement stroke of $x + h$, whose site-average is the myosin working stroke h. For the trapped filament and rigid links, these displacements are reduced by a factor $k/(k + \kappa_t)$; correction factors for the case of compliant links are derived in Appendix D.

aim to detect individual events, which is the focus of the present paper. However, there is no general agreement on the size of the working stroke: substantially different values have been reported (e.g., Guilford et al., 1997; Ishijima et al., 1998; Veigel et al., 1998), which could arise from differences in experimental method or analytical technique.

The analytical approaches adopted so far do not provide in any sense a "best fit" to the data. In this paper we describe a method that optimally fits data to a simple kinetic model using a Hidden-Markov (HM) algorithm; similar algorithms have been used to detect gating events in single-ion-channel recordings (Colquhoun, 1998; Ball and Sansom, 1989; Chung et al., 1990, 1991; Fredkin and Rice, 1992; Horn and Lange, 1983). For event detection, HM algorithms are more powerful than single-edge algorithms because they generate optimized model parameters with respect to the whole time series (Appendix A) and the user need not set a threshold level that may influence the number and kind of detected events.

A single myosin molecule makes stochastic transitions between states characterized by differences in biochemistry, conformation, or attachment to actin. These states, or a subset of them, must be inferred from a noisy displacement record and are in that sense "hidden." The "Variance-Hidden-Markov" method presented in this paper operates on a variance record that monitors attachment events only.

The variance-HM method requires only the simplest possible assumptions that will account for the experimental data: 1) the displacement record consists of a sequence of levels (the "signal") plus stationary random noise; 2) myosin binds intermittently to F-actin in a two-state model, either bound to or free from actin; and 3) when myosin binds to actin the variance of bead displacements decreases because the longitudinal stiffness of the dumbbell-myosin system exceeds the stiffness of the dumbbell alone. The

parameters of this model are the rate constants f and g for myosin binding and detachment, and the free and bound variance levels V_1 and V_2. This information enables the probability of each variance point and attachment state (free or bound) in the variance record to be calculated from the variance and state of the preceding point; the likelihood of the entire sequence is obtained by compounding these conditional probabilities, given a suitable assumption about the initial state. The likelihood is then maximized with respect to the parameters f, g, V_1, V_2. A running window of W data points is used to calculate variance as a function of time, and a figure of merit related to the likelihood indicates the optimum value of W.

Readers content with the above explanation may wish to omit the mathematical description in the next section (The Variance-Hidden Markov Method), and turn to the following section (Experimental Constraints). There follows an account of tests with Monte Carlo-simulated input, in which the actin dumbbell is assumed to be internally rigid so that the beads make identical displacements. This assumption is relaxed in the next section (Two-Channel Data) for the case where actin-bead links are compliant; then the displacement records of the two beads differ and the observed working stroke must be corrected for the compliant links. We also describe a method of correcting for a slowly drifting baseline. Results are presented for two-channel Monte Carlo simulations and experimental data for rabbit myosin-S1.

The variance-HM method will not detect actomyosin events after binding that generate a step change in displacement without increasing the stiffness of the myosin "crossbridge." Events linked to force generation in muscle (Huxley and Simmons, 1971) or force holding in myosin I (Veigel et al., 1999) may be of this kind. These limitations can be overcome by other versions of the HM method not considered in this paper. The strengths and limitations of the method are summarized in the penultimate section (Using the Method). Computational aspects of the variance-HM method and computer programs are presented in supplementary material.

In this paper the unit of frequency is assumed to be radians per second, the only exception being for data collection where the inverse of the sampling interval is traditionally expressed in Hertz. The meanings of most mathematical symbols in the text are summarized in Table 1.

THE VARIANCE-HIDDEN-MARKOV METHOD

We describe the basis of the HM method for a single record $(v_1, \ldots, v_N)$ of variance values at equal intervals Λ in time. In the process we consider how this variance record can be constructed from a record of bead displacements.

The variance record is assumed to derive from a model with only two states, specifying attachment of myosin to actin (1 = "off," 2 = "on"). Let $A_{ij}(v, v')$ be the probability of a transition from a recorded point with variance v and

TABLE 1 Digest of mathematical symbols used in the main text

$A_{ij}(v, v')$ = transition rate from state i, variance v to state j, variance v';
a_{ij} = transition rate from state i to state j (i, j = 1 or 2);
c = actin monomer spacing;
F = figure of merit = **L**/N + εW ($\varepsilon \sim$ 0.0015–0.003);
f, g = myosin-actin binding/detachment rates;
h = unloaded myosin working stroke;
i = myosin-actin attachment state (1 = free, 2 = bound);
J_i = number of detected attachment events into state i;
J = total number of detected attachment events $\equiv J_1 + J_2$;
k = myosin stiffness;
k_B = Boltzmann's constant;
L = $\log_{10}(P)$ (P = likelihood);
N = no. of entries in the variance record;
R = signal-to-noise ratio;
S_i = SD of all displacement levels in state i;
T = absolute temperature;
T = duration of displacement record;
$u(t)$ = bead displacement at time t;
$\bar{U}_i$ = mean (event-averaged) displacement in state i;
$\bar{U}$ = uncorrected working stroke $\equiv \bar{U}_2 - \bar{U}_1$;
V_i = mean (time-averaged) variance level in state i;
V_{iLR} = mean covariance level in state i;
W = no. of data points in the running-variance window;
β = damping constant of both beads;
Δ = time interval for sampled displacements;
$\delta_{i,j}$ = Kronecker delta function = 1 ($i = j$), 0 ($i \neq j$);
λ_t = corner frequency of free dumbbell;
λ = corner frequency of bound dumbbell;
Λ = time interval for variance record = $(W/2)\Delta \equiv T/N$;
κ_t = combined stiffness of both traps;
κ_L, κ_R = stiffnesses of left, right bead-filament linkages;
$\tilde{\kappa}_L, \tilde{\kappa}_R$ = stiffnesses of left, right links in series with their traps;
$\tilde{\kappa}$ = combined stiffness of filament-bead system = $\tilde{\kappa}_L + \tilde{\kappa}_R$
$\varphi_i(v)$ = distribution of variance v in state i;
σ_i = predicted SD of displacements in state i;
ω_c = Nyquist sampling frequency (variance record);

state i to variance variance v' and state j at the next time-point, where i, j = 1 or 2. By setting these variances in turn to the observed values (v_1, v_2), (v_2, v_3) and so on, multiplying the corresponding chain of transition probabilities $A_{i1i2}(v_1, v_2)$, $A_{i2i3}(v_2, v_3)$, . . . , and summing all attachment states $(i_1, i_2, \ldots, i_N)$ over values (1, 2), one obtains the probability P (or "likelihood") that the chosen transition matrix fits the entire variance record. The parameters of the transition matrix may then be varied to maximize the value of P. An efficient method of implementation is due to Baum (1972) and is described in general terms in Appendix A, which applies here if $x \equiv v$.

In this paper we use the classical Hidden-Markov method (Rabiner, 1989) where the transition matrix is independent of the initial variance v, and can therefore be expressed in the form

$$A_{ij}(v, v') = a_{ij}\varphi_j(v') \quad (i, j = 1, 2). \tag{1}$$

where $\varphi_j(v)$ is normalized to unity. In this case, variance noise is regenerated from a state-dependent distribution

$\varphi_j(v)$ at each sampled time, and is uncorrelated over the sampling interval Λ. The "states-only" transition matrix a_{ij} can be expressed in terms of notional binding and detachment rates f and g as

$$a_{12} = \frac{f}{f + g}\left(1 - e^{-(f+g)\Lambda}\right),$$

$$a_{21} = \frac{g}{f + g}\left(1 - e^{-(f+g)\Lambda}\right), \tag{2}$$

with a_{11} and a_{22} determined from the sum rule over final states (Eq. A7b). These expressions follow from the corresponding rate equations for $a_{ij}(\Lambda)$ as a function of Λ.

The distributions $\varphi_j(v)$ are determined by how the variance record is constructed, for example by a uniform average over a forward window of W sample points:

$$v(t_k) = \frac{1}{W}\sum_{j=1}^{W}(u(t_{k+j}) - \bar{u}(t_k))^2, \quad \bar{u}(t_k) = \frac{1}{W}\sum_{j=1}^{W}u(t_{k+j}) \tag{3}$$

where bead displacements $u(t)$ are collected at times $t_k = k\Delta$ (k = 1, 2, . . .). In general, variance values need not be constructed at every time point of the displacement record, because any correlation between neighboring points in the displacement record will then appear in the variance record, which conflicts with the assumption leading to Eq. 1. We show below that maximum information compatible with minimal correlations can be achieved if variances are generated or resampled at every $W/2$th point of the displacement record, so that the reduced sampling time Λ is $(W/2)\Delta$.

If the displacements $u(t_k)$ are independent Gaussian variates with the same attachment state, mean and variance σ_i (Gaussian white noise), then $Wv(t_k)/\sigma_i^2$ is a chi-square function with $W - 1$ degrees of freedom (Weatherburn, 1968), so

$$\varphi_i(v) = \frac{W}{2\sigma_i^2}\frac{y_i^{(W-3)/2}e^{-y_i}}{\Gamma((W-1)/2)}\left(y_i = \frac{Wv}{2\sigma_i^2}\right). \tag{4}$$

Hence the mean and variance of v in attachment state i are

$$V_i = (1 - W^{-1})\sigma_i^2, \quad S_i^2 = (1 - W^{-1})\frac{2\sigma_i^4}{W}. \tag{5}$$

where V_i is the true level of the running-variance record in that state. The standard deviation S_i from this level is such that $S_i/V_i \sim \sqrt{2/W}$ for $W \gg 1$. Hence the running-variance record can be regarded as the sum of signal and noise components; the signal switches between values V_1, V_2 and the noise amplitudes S_1, S_2 are state-dependent. The quantity

$$R \equiv \frac{V_1 - V_2}{S_1 + S_2} \approx \sqrt{\frac{W}{2}}\frac{V_1 - V_2}{V_1 + V_2} \tag{6}$$

may be regarded as a signal-to-noise ratio. Similarly, the autocorrelation function

$$C_i(\tau_n \equiv n\Delta) \equiv \overline{(v(t_k) - V_i)(v(t_{k+n}) - V_i)}$$

$$= \begin{cases} S_i^2 \left\{ 1 - \left(1 - \dfrac{n}{W^2} \right) \dfrac{n}{W-1} \right\}, & (n \leq W) \\ 0, & (n > W). \end{cases}$$

$$(7)$$

for variance fluctuations in state i arises from Gaussian white displacement noise; as expected, fluctuations are uncorrelated when $n > W$. The correlation time, defined as the time to half-maximum correlation, asymptotes to $W/2$ sample points when $W \gg 1$. Choosing $\Lambda = (W/2)\Delta$, the correlation time of variance fluctuations, gives an acceptable trade-off between time resolution and unwanted correlations. This choice can also be justified in the frequency domain. In the limit $W \gg 1$, the corresponding power spectrum with respect to circular frequency ω is proportional to $\sin^2(\omega\Lambda)/(\omega\Lambda)^2$, and 90% of the total spectrum lies within the first lobe where $|\omega| < \pi/\Lambda$. As π/Λ is the Nyquist critical frequency for the sampling interval Λ, aliasing effects from higher frequency components should be minimal (Press et al., 1992). Note also that, for a given value of Δ, a large averaging window gives a bigger signal-to-noise ratio, but degrades the time resolution ($\pm\Lambda/2$) defined by the variance record.

Events are detected by using Baum's forward-backward algorithm (Appendix A) to generate the probabilities p_1, $p_2, \ldots, p_N$ at each point of the variance record that myosin is bound, which indicate the more-probable attachment state of myosin. Then the parameters f, g, V_1, V_2 can be re-estimated by Baum's method (Appendix A), enabling an iterative approach to an optimum parameter set for which the likelihood has a global maximum. The maximum with respect to f, g is typically broader than with respect to V_1, V_2 (Fig. 2). In the absence of round-off errors, Baum's theorem (Appendix B) guarantees that the iterated likelihood over a cycle of re-estimation will not decrease.

Averaged quantities calculated after event detection are accompanied by finite sampling errors. For example, event-averaged displacement levels $\bar{U}_1$, $\bar{U}_2$ with myosin free or bound and their standard deviations S_1, S_2 are calculated by the formulae

$$\overline{U}_i = \frac{1}{J_i} \sum_{j=1}^{J} U(j)\delta_{i(j),i}, \quad \overline{U_i^2} = \frac{1}{J_i} \sum_{j=1}^{J} U(j)^2 \delta_{i(j),i} \quad (8a)$$

$$S_i^2 = \overline{U_i^2} - (\overline{U}_i)^2 + \frac{1}{J_i} \sum_{j=1}^{J} \frac{S(j)^2}{n(j)} \delta_{i(j),i} \quad (8b)$$
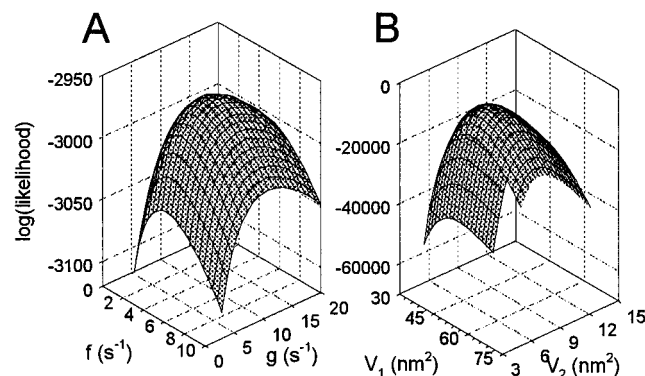

FIGURE 2 Sensitivity of log-likelihood **L** to changes in (A) rate constants f and g and (B) variance levels from a Monte Carlo simulation. Each surface was calculated using the optimum values of the other two parameters ($V_1 = 53.84$, $V_2 = 6.683$ nm^2 in A and $f = 4.548$, $g = 10.07$ s^{-1} in B). Maximum likelihood with respect to f and g may be achieved by Baum's re-estimation method (Appendix B), nested within a cycle of re-estimation for variance levels.

for J_1 detected detachments and J_2 bindings (Weatherburn, 1968). $U(j)$ and $S(j)$ are the mean and standard deviation of displacement in the jth attachment period, containing $n(j)$ sample points. Hence $E_i = S_i/\sqrt{J_i}$ is the standard error in the mean level in state i. However, these errors do not include the effects of mistakes in event detection.

EXPERIMENTAL CONSTRAINTS

Under what conditions can the variance-HM method be applied to displacement data from the double-bead trap experiment (Fig. 1)? There are restrictions relating to the time scales of events and data collection, and conditions for a high signal-to-noise ratio to ensure reliable event detection. In this section we assume that the actin-bead dumbbell is sensibly rigid so that the displacement records from the two beads are the same. Formulae presented in this section apply also to the average of these two records when the compliances of the two actin-bead links are identical (Smith, 1998b).

In the previous section, displacement noise was approximated by Gaussian white noise to estimate the statistical properties of the derived variance record. This assumption is incorrect, since trap force and viscous damping on each bead act in combination to limit the frequency response of longitudinal bead displacements below the "corner" frequencies

$$\lambda_t = \kappa_t/\beta, \quad \lambda = (k + \kappa_t)/\beta \quad (9)$$

(in radians/s) with myosin respectively free or bound (Svoboda and Block, 1994). Here κ_t is the combined stiffness of the traps, β the damping constant of both beads, and k the myosin stiffness, corrected if necessary for compliant actin-bead links (Fig. 1). These frequencies are rates of mechan-

ical damping, or regeneration of an equilibrium distribution of Brownian fluctuations. The variances of displacement noise with myosin free or bound are determined by equipartition of energy, namely

$$\sigma_1^2 = \frac{k_{\mathrm{B}}\mathrm{T}}{\kappa_{\mathrm{t}}}, \quad \sigma_2^2 = \frac{k_{\mathrm{B}}\mathrm{T}}{k + \kappa_{\mathrm{t}}} \tag{10}$$

where k_{B} is Boltzmann's constant and T is absolute temperature). When $W \gg 1$, the variance levels V_1, V_2 are almost equal to σ_1, σ_2 respectively (Eq. 5).

For displacements at intervals of Δ, the approximation of Gaussian white noise is correct if $\lambda_{\mathrm{t}} \gg \pi/\Delta$. For the variance-HM method, this condition can be relaxed because the variance record has a larger sampling time Λ. In the last section we showed that the choice $\Lambda = (W/2)\Delta$ is sufficient to reduce correlation between successive variance points to an acceptably low level. In practice, Brownian equilibrium between variance points is sufficiently well established if λ_{t} exceeds π/Λ by a factor of three.

A high signal-to-noise ratio in the variance record is achieved if 1) in bound periods, the mean variance drops to a small fraction of its free level and 2) variance fluctuations in both free and bound periods are low. The former is accomplished by using weak traps relative to the stiffness of myosin and links in series, and by meeting the conditions for Brownian equilibrium in the variance record. In free and bound periods, variance fluctuations can be reduced to a small fraction of the variance level by increasing the width W of the averaging window (Eq. 5 and below).

The variance-HM method presented here requires the following working conditions: 1) to ensure uncorrelated variance values, Brownian equilibrium should be achieved within the time interval Λ between sampled points in the variance record; 2) in general, each period of attachment should last for many such sample points. Simulations suggest that periods with lifetimes below 4Λ may be missed; 3) for reliable event detection, the variance levels V_1, V_2 in free and bound periods should be well separated relative to fluctuations, as expressed by the inequalities

$$\lambda_{\mathrm{t}} \gg \frac{2\pi}{W\Delta} \gg f, g. \tag{11}$$

$$\frac{V_1}{V_2} \approx \frac{k + \kappa_{\mathrm{t}}}{\kappa_{\mathrm{t}}} > 1 \tag{12a}$$

$$R \equiv \sqrt{\frac{W}{2}} \frac{V_1 - V_2}{V_1 + V_2} \approx \sqrt{\frac{W}{2}} \frac{k}{k + 2\kappa_{\mathrm{t}}} > 1. \tag{12b}$$

where R is the signal-to-noise ratio of Eq. 6. For a given value of R, a bigger window is required when the variance ratio in Eq. 12a is reduced, say by weaker links. [Note the equivalence of the units of rate constants (s^{-1}) and frequency (rads/s) implied by Eq. 11. The characteristic frequency of a relaxation process $\exp(-\lambda t)$ is λ, since its

power spectrum with respect to circular frequency ω is proportional to $(\omega^2 + \lambda^2)^{-1}$. This form motivates (11) when ω is the Nyquist sampling frequency $\pi/(W\Delta/2)$ in the variance record].

Estimated values for trap experiments are $\kappa_{\mathrm{t}} = 0.04$ pN/nm, $\beta = 2 \times 10^{-8}$ N·s/m for two 1-μm diameter beads in aqueous solution (Stokes' law, viscosity = 0.001 Kg/m·s), hence $\lambda_{\mathrm{t}} = 2000$ s^{-1}, $k = 0.1$–1.0 pN/nm with or without link compliance (Veigel et al., 1998; Huxley and Tideswell, 1996), $\Delta = 10^{-4}$ s and $f = 5$ s^{-1}, $g = 10$ s^{-1}. The binding rate f is very dependent on dumbbell positioning, but the rate of detachment should reflect rates of ATP-induced detachment measured by other means, for example $g \approx 2$–$3 \times 10^6[\mathrm{ATP}]$ s^{-1} below 100 μM (Lymn and Taylor, 1971). These values satisfy conditions (Eq. 11) if $W = 100$, giving $\Lambda = 5$ ms. The quoted stiffnesses give $V_1/V_2 = 3.5$–26 where values above 5 are desirable, and $R = 3.9$–6.5.

Additional assumptions, extraneous to the variance-HM method or any event detector using a running-variance record, are required to extract the myosin working stroke from displacement data. The method proposed by Molloy et al. (1995) assumes that all sites on F-actin are equally accessible, so that the distribution of bound displacement levels is the Boltzmann distribution of free dumbbell displacements displaced by an amount

$$\bar{U} = \frac{k}{k + \kappa_{\mathrm{t}}} h \tag{13}$$

(Appendix C) where h is the unloaded myosin working stroke. For equally compliant links, this equation is correct if link compliance is absorbed in k (Smith, 1998b; Veigel et al., 1998). General correction formulae for compliant links are derived in the section on two-channel data.

TESTS WITH MONTE CARLO SIMULATIONS

The variance-HM method as an event detector has been extensively tested with Monte Carlo simulations, made using the Langevin equation method (Smith, 1998b). All simulations were made with a simplified actomyosin cycle as follows: 1) myosin-S1 can bind to monomers of any orientation on the actin filament (Molloy et al., 1995); 2) the rate of binding is a symmetric (Gaussian) function of the myosin-actin site separation x in Fig. 1; and 3) binding is immediately followed by a force-generating transition that moves the neck-rod junction by the "throw distance" or working stroke h. These features produce quantized displacement levels separated by the smallest monomer spacing c on the double-helix that allows the myosin to bind (Lorenz et al., 1995). The time step for integrating the Langevin equation should be within mechanical damping times (under 1 μs), while displacement output was collected
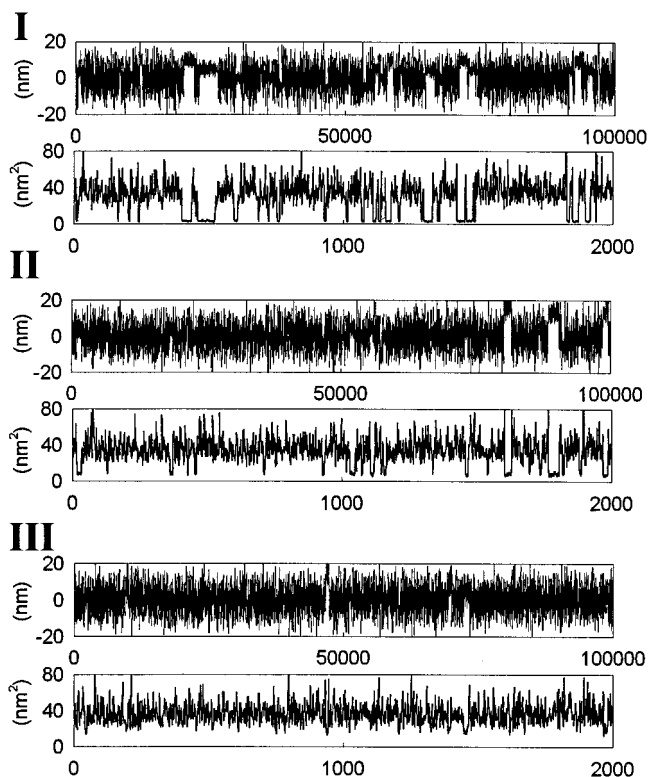
FIGURE 3 Samples of Monte Carlo-simulated displacement data (one-tenth of a 100-s record sampled at 10 KHz) and their running variance records (with $W = 100$, sampled at 200 Hz) for $h = 5.5$ nm and variance ratios of 11, 5, and 2 (I, II, III, respectively). Each simulation is made for a combined trap stiffness of 0.1 pN/nm, and values $k = 1.0$, 0.4, and 0.1 pN/nm, respectively of effective myosin stiffness. For compliant but equal actin-bead links, k can be interpreted as the stiffness of myosin in series with the links in parallel, provided both bead displacements are averaged (Smith, 1998b). Simulation kinetics determine the Brownian-averaged rates of binding $f_{eq} = 4.57$ s^{-1} (Eq. C7) and detachment $g_{eq} = 12.7$, 21.1, and 33.1 s^{-1}, respectively. The corresponding duty ratios (0.26, 0.18, 0.12) are typical of experiments at 5 μM ATP. Displacements were simulated from a Langevin equation with randomly generated Brownian forces and attachment events. For clarity, only every 10th displacement point is plotted.

at 10 KHz ($\Delta = 10^{-4}$ s). Simulations were made with $c = 5.5$ nm and $h = 5.5$ nm or 11 nm.

This section describes results from a simulation program with one output channel. Simulations incorporating various experimental complications, including links of different stiffness, are analyzed in the next section.

To establish a benchmark, we present results from a simulation of 10^6 points with $h = 5.5$ nm, a high variance ratio ($V_1/V_2 = 11$), and equivalent rate constants $f_{eq} = 4.6$ s^{-1}, $g_{eq} = 12.7$ s^{-1} (Eqs. C7, C8), a fragment of which is shown in Fig. 3 I. With $W = 100$, the core program detected 567 attachment events of 641, and produced re-estimated values $f = 4.0$ s^{-1}, $g = 10.0$ s^{-1}, and $V_1/V_2 = 9.6$ convergent to within 0.1%. The detected rate constants are lower than predicted, either because short-lived events in the sim-

ulation have not been detected or are not present in the variance record. During re-estimation, log(likelihood) $\mathbf{L} = \log P$ rose monotonically to a maximum. Fig. 4 summarizes input and output information from the program suite, including a fragment of the running-variance record (A) and its distribution (B), "restoration" of the displacement record after event detection (C) as a sequence of flat levels (mean displacement in each detected free or bound period), and the frequencies of binned free and bound levels (D) and their lifetimes (F). Note the following features: 1) the spread of free levels is confined to a few nanometers and should be symmetric about the base line unless contaminated by undetected bound periods; 2) the distribution of bound levels reflects the Gaussian distribution of free dumbbell displacements with standard deviation σ_1 (Appendix C), displaced to a mean value slightly less than the working stroke h; 3) detected bound levels reflect the quantization of levels produced by the simulation from discrete binding sites on the filament (Fig. 4 E); and 4) lifetimes of free and bound levels are exponentially distributed with the detected decay rates f and g, except that some short-lived levels appear to be missing (Fig. 4 F).

The detected rate constants are related to the number of detected events, since the steady-state flux through the attachment cycle is $fg/(f + g)$. Assuming steady-state conditions, the expected numbers of detachment/binding events over time T are

$$J_1 = J_2 = \frac{fg}{f + g} T. \tag{14}$$

If $J_1 > 100$, this equation predicts the number of detected events to the nearest integer.

Indicators of performance

As a detector of events, the variance-HM method can be characterized at various levels. For simulations, the times of binding and detachment events in the displacement record are known, so the incidence of missed events and false detections can be tabulated. These numbers are useful in comparing results from variance records with different windows from the same displacement record, and more generally in testing the method over a range of values of simulation parameters.

For experimental data, other indicators are required. From assorted output variables, the most useful guides seem to be the maximized likelihood $\mathbf{L}$, mean bound level $\bar{U}$ (the raw working stroke), the number of detected events J, and detachment rate g as a function of window width. Information revealed by level and lifetime histograms can also be useful: 1) the distribution of free levels may reveal a loss of long-term stability in the positions of the traps or the microscope stage; 2) the distribution of bound levels may reveal departures from Gaussian behavior; and 3) the dis-
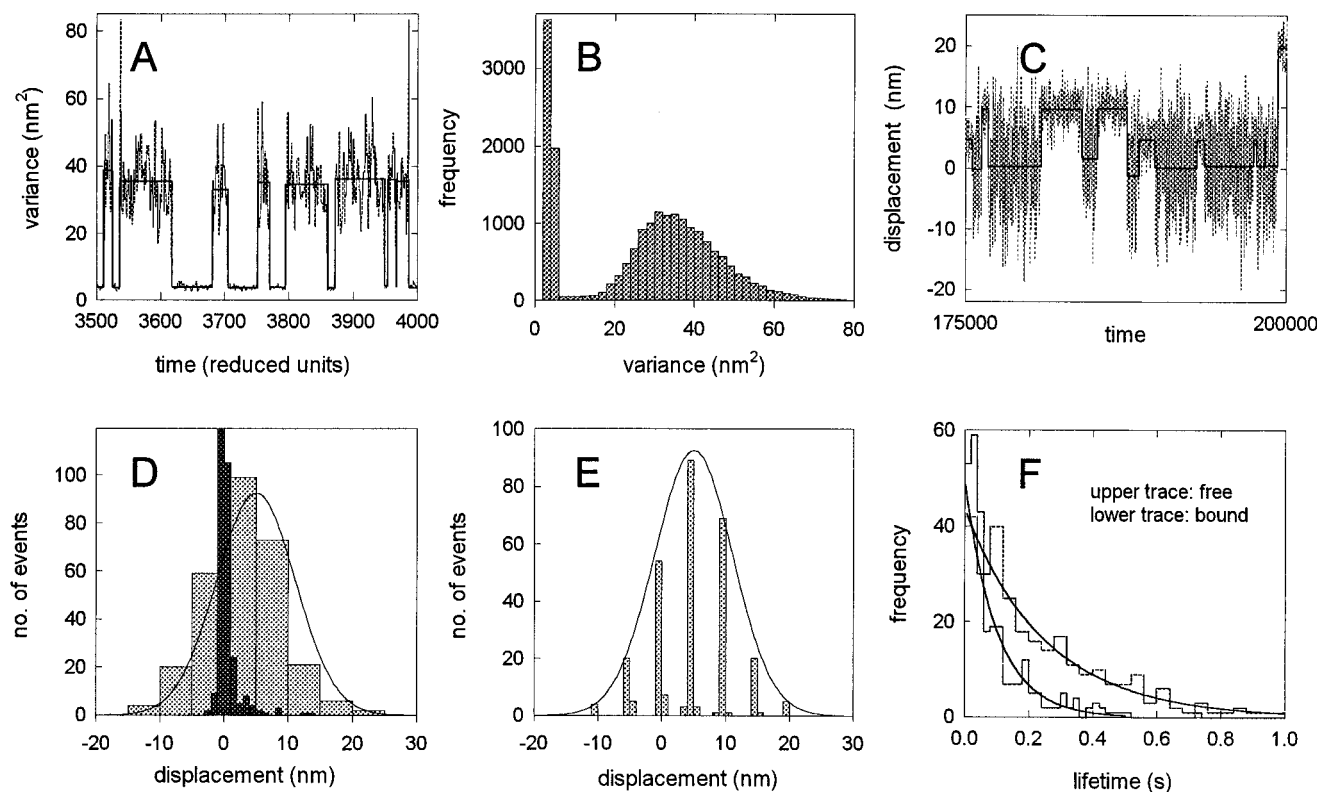
FIGURE 4   The variance-HM method applied to displacements simulated in Fig. 3 I, showing (*A*) a fragment of the running-variance record, overlaid by the sequence of detected variance levels; (*B*) the bimodal distribution of variances in the record, reflecting the predicted mean variance levels $V_1 = 40.0$, $V_2 = 3.64$ nm^2 in free and bound periods, with fractional widths approximately equal to $\sqrt{2/W} = 0.14$ from Eq. 5. (*C*) A fragment of the displacement record, "restored" by the sequence of detected displacement levels. Short-lived features of the simulation are either not detected or not present in the variance record. (*D*) Frequencies of detected detachments (*dark shading*) and bindings (*light shading*) selected by the following displacement level, using 1-nm and 5-nm bins, respectively, and the relation of the latter to a Gaussian distribution matching the sampled mean (5.00 nm), standard deviation (6.13 nm), and total number of detected bindings (284). (*E*) The same distribution of binding events, binned at one-fifth of the expected quantization spacing of 5.0 nm. (*F*) Distributions of the lifetimes of free and bound levels, and their relation to Poisson distributions of degree zero with mean lifetimes $1/f_{opt} = 0.25$ s and $1/g_{opt} = 0.10$ s.

tribution of lifetimes may indicate that some short-lived events have been missed.

With simulated input, detected binding or detachment events can be classed as true or false by matching them in time to known input events. A similar procedure can define undetected events in the input record. Input and output events will be said to be matched if their time difference is within one window width (a time error of $\pm W\Delta$), although variance values are stored at half this interval. The construction of such comparison algorithms is not straightforward and is discussed in the supplement.

Event detection can also be judged by the value of the maximized likelihood, or log(likelihood)/point **L**/N, which is approximately independent of the length N of the variance record for large N. Table 2 lists values of **L**/N and other indicators as a function of window width for simulations I–III of Fig. 3 with different variance ratios. The number of detected events and the apparent attachment rates are decreasing functions of W. In simulation I, $\bar{U}$ tends to the correct value at small W because no false detections were

introduced. In simulations II and III, false events are present so $\bar{U}$ moves toward its correct value as W is increased and false detections are reduced in number. Thus the optimum window width is determined by the nature of the data.

Although maximum likelihood indicates the best parameters for a given variance record, it does not indicate the best window size for a given displacement record. A similar difficulty arises in attempting to compare maximum-likelihood values from models with different numbers of parameters, and empirical corrections have been suggested as a way of penalizing over-specified stochastic models (Ball and Sansom, 1989). Here we wish to characterize outputs from a fixed stochastic model when information in the variance input is reduced, suggesting an additive bonus proportional to W. For these purposes, an empirical figure of merit

$$\mathbf{F} = \frac{\mathbf{L}_{max}}{N} + \varepsilon W \qquad (15)$$

TABLE 2 Performance of the variance-HM detector as a function of window width *W* and the free-to-bound variance ratio V_1/V_2, using simulations I–III of Fig. 3

V_1/V_2	W	$\mathbf{L}_{max}/N$	f_{opt} (s⁻¹)	g_{opt} (s⁻¹)	$\bar{U}$ (nm)	J	$J^{(T)}$	$J^{(M)}$	$J^{(F)}$
11 (I) (641)	50	−1.5801	4.29	10.17	4.69	601	601	39	0
	100	−1.6712	3.97	9.97	4.49	567	566	73	0
	150	−1.8043	3.57	9.45	4.53	517	517	120	0
	200	−2.0380	3.37	9.38	4.43	495	495	140	0
	250	−2.3334	2.99	8.92	4.29	447	447	186	0
5 (II) (707)	50	−1.7957	6.14	27.61	2.75	990	652	54	338
	100	−1.8535	3.72	18.82	3.94	620	600	104	20
	150	−1.8985	3.26	17.84	4.01	549	548	154	2
	200	−2.0095	3.00	17.55	3.98	510	506	188	1
	250	−2.1144	2.65	16.61	4.15	456	455	238	0
2 (III) (714)	100	−1.7506	22.1	61.3	0.44	3184	539	170	2632
	150	−1.7663	10.34	43.8	0.76	1635	500	207	1131
	200	−1.7837	6.51	34.5	1.03	1063	454	251	599
	250	−1.8040	4.52	28.9	1.51	771	411	296	357
	300	−1.8372	3.39	25.6	1.69	597	364	341	228

The total number of events is given in parentheses. J is the number of detected events; superscripts T, M, F denote true, false, and missing events as defined in the main text. Other symbols are defined in Table 1. The expected rate constants are $f_{eq} = 4.57$ s⁻¹ throughout and $g_{eq} = 12.7$, 21.1, and 33.1 s⁻¹ (Appendix D), also mean bound levels $\bar{U} = 5.0$, 4.4, and 2.75 from Eq. 13. With simulation I there are no false detections, and small windows detect more true events and give the best results. With simulations II and III, false events are detected. The incidence of false positives is reduced by increasing the window size, thereby increasing the estimated mean bound level. For simulation III, the method fails as the number of false detections remains unacceptably high as W is increased. Note that the optimum value of W is not indicated by maximum likelihood.

is proposed. The coefficient ε should be chosen so that **F** and the raw working stroke $\bar{U}$ are maximized at the same value of W, since simulations show that the maximum value of $\bar{U}$ is closest to the value predicted by Eq. 13.

Fig. 5 summarizes the behavior of four indicators, namely (*A*) the figure of merit, (*B*) raw working stroke, (*C*) number of detected events, and (*D*) optimized detachment rate, as a function of window width for the simulations in Table 2. The desired correlation between **F** and $\bar{U}$ is achieved in all four simulations with the same value (0.0015) of ε, so that

the window size for maximum **F** reflects both the accuracy of the raw working stroke and the quality of the data (as in Table 2). Results with this window should be accepted unless contraindicated by *C* and *D*, as described in the figure caption. There is no universal "best" value of ε, simulations made with a lower trap stiffness (0.067 pN/nm) and $h = 11$ nm behaved similarly but required $\varepsilon = 0.0028$. Thus the figure of merit can be regarded as an empirical adjunct to indicators *B*, *C*, and *D*, which should be valid for data records obtained under the same conditions.
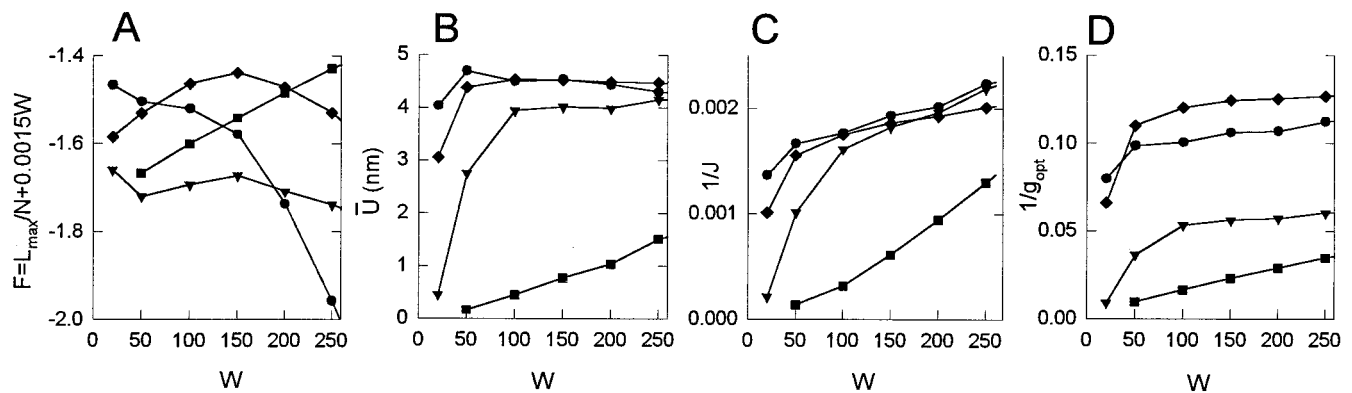

FIGURE 5 Performance indicators for variance-Hidden-Markov runs at different window sizes W, for simulations I–III of Fig. 3 (●, ▼, ■) and from one channel of a two-bead simulation (◆). (*A*) The figure of merit **F**; (*B*) mean bound displacement level $\bar{U}$; (*C*) the reciprocal of the number of detected events J; and (*D*) the mean lifetime $1/g_{opt}$, where g_{opt} is the optimized detachment rate. For simulations I–III, values 5.0, 4.4, and 2.75 nm, respectively, are predicted for $\bar{U}$ and values 12.7, 21.1, 33.1 s⁻¹ for g. The best value of W is indicated by a maximum in the figure of merit, for which $\bar{U}$ is close to its predicted value. In all plots, false detections become significant for $W < 50$, producing a spurious drop in $\bar{U}$ and high values of J and g_{opt}. As W increases above 50, J and g_{opt} fall more slowly as genuine short-lived events are progressively removed from the running-variance record. When false detections are absent, the mean lifetime increases linearly with W and can be fitted to the equation $1/g_{opt}(W) = 1/g + \alpha(W\Delta/2)$, in which the second term is the shortest detected lifetime. All events present in the variance record are detected if $\alpha = 1$. Data for simulation I can fit such a straight line for $W > 100$.
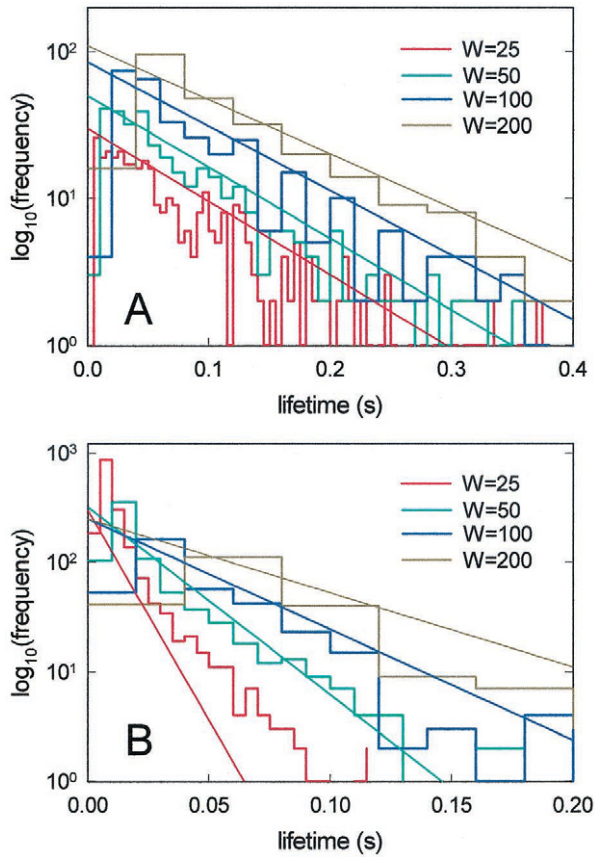
FIGURE 6 Log-linear histograms of the lifetimes of bound levels as a function of window size W in the running-variance record can reveal and characterize missing and false events. (*A*) For good data ($V_1/V_2 = 11$), short-lived bound events are missed as W is increased, but longer-lived states fit a Poisson distribution with the detected detachment rate (*straight lines* with slopes $g_{opt} = 11.5$, 11.2, 10.1, and 8.5 s^{-1}, respectively). The reciprocals of these values fit the straight-line function described under Fig. 5 with $\alpha = 3.6$ and $g = 12.3$ s^{-1} (simulation value $= 14.9$ s^{-1}), indicating that false events are virtually absent. (*B*) For bad data ($V_1/V_2 = 2$), many short-lived false events appear with a small window, so the detected detachment rate does not describe the distribution of longer-lived events; as W increases the peak at short lifetimes is replaced by a hole reflecting missing events and the detachment rate drops more rapidly ($g_{opt} = 88.4$, 39.5, 23.2, and 15.6 s^{-1}). For reasonable bin statistics, bound-state lifetimes are binned at four times the sampling time $\Lambda = 4(W\Delta/2)$ in the running variance record.

Missed events come in pairs, from undetected short-lived levels. Their number can be estimated by comparing lifetime histograms at different window widths, which impose a threshold time $\Lambda = W\Delta/2$. In Fig. 6 *A* the lifetimes of missed levels are $<4\Lambda$ and the detection of longer-lived levels is virtually unchanged; this behavior is also a positive indicator. False events may appear as a sharp increase in J and g_{opt} as W is reduced, where g_{opt} does not describe the distribution of longer-lived events (Fig. 6 *B*). The apparent working stroke is also reduced, as false events arising from noise are symmetrically distributed about the baseline.

Simulations also reveal the amount of variability in detected output from Monte Carlo simulations with different random-number seeds but the same physical parameters. In case I of Fig. 3 ($V_1/V_2 = 11$), 10 such Monte Carlo runs of 10^6 points each generated raw working strokes over a range of 0.49 nm (mean $\pm$ SD $= 4.89 \pm 0.15$ nm) near the true value of 5.0 nm, with 250–300 detected binding events per record. These variations mostly lie within the estimated standard error $E \approx 0.4$ nm for each record where $E^2 = (S_1^2 + S_2^2)/J_2$, so variations between runs are generated by the same fluctuations that generate events within each run. However, histograms of bound levels from individual simulations do show significant differences. Repeated simulations of types *B* and *C* behave similarly, except that false detections were present at the window width used ($W = 100$) and the raw working strokes were lower.

Output may be combined in ways determined by the kind of averaging used in each record and statistical assumptions about variability between records. Displacements levels are "event-averaged," with a weight of one for each attachment period, so the average displacement over n records should be calculated by event-averaging the combined record. Let $U^{(r)}$ and $S^{(r)}$ be the mean and standard deviation of displacements in a given state in the rth record, with $J^{(r)}$ events into the state; the corresponding standard error is $E^{(r)} = S^{(r)}/\sqrt{J^{(r)}}$. If $p^{(r)}$ is the fraction of events in the rth record, the mean displacement is

$$\langle U \rangle = \sum_{r=1}^{n} p^{(r)} U^{(r)} \quad \left(p^{(r)} = \frac{J^{(r)}}{J_{tot}}, J_{tot} = \sum_{r=1}^{n} J^{(r)} \right). \quad (16)$$

The standard deviation S is given by combining variances between and within records:

$$S^2 = \langle U^2 \rangle - \langle U \rangle^2 + \sum_{r=1}^{n} p^{(r)} S^{(r)2}. \quad (17)$$

The standard error in the combined result is $E = S/\sqrt{J_{tot}}$, so

$$E^2 = \frac{\langle U^2 \rangle - \langle U \rangle^2}{J_{tot}} + \sum_{r=1}^{n} p^{(r)2} E^{(r)2} \quad (18)$$

(Weatherburn, 1968). However, the derivation of Eqs. 17 and 18 assumes that there are no systematic variations between records. If such variations exist, the standard error would be obtained by replacing J_{tot} by the number of records n in Eq. 18. These formulae can detect systematic changes in working strokes from different records.

The variety of simulations

Using the same indicators (the figure of merit, raw working stroke, number of detected events, and optimized detach-
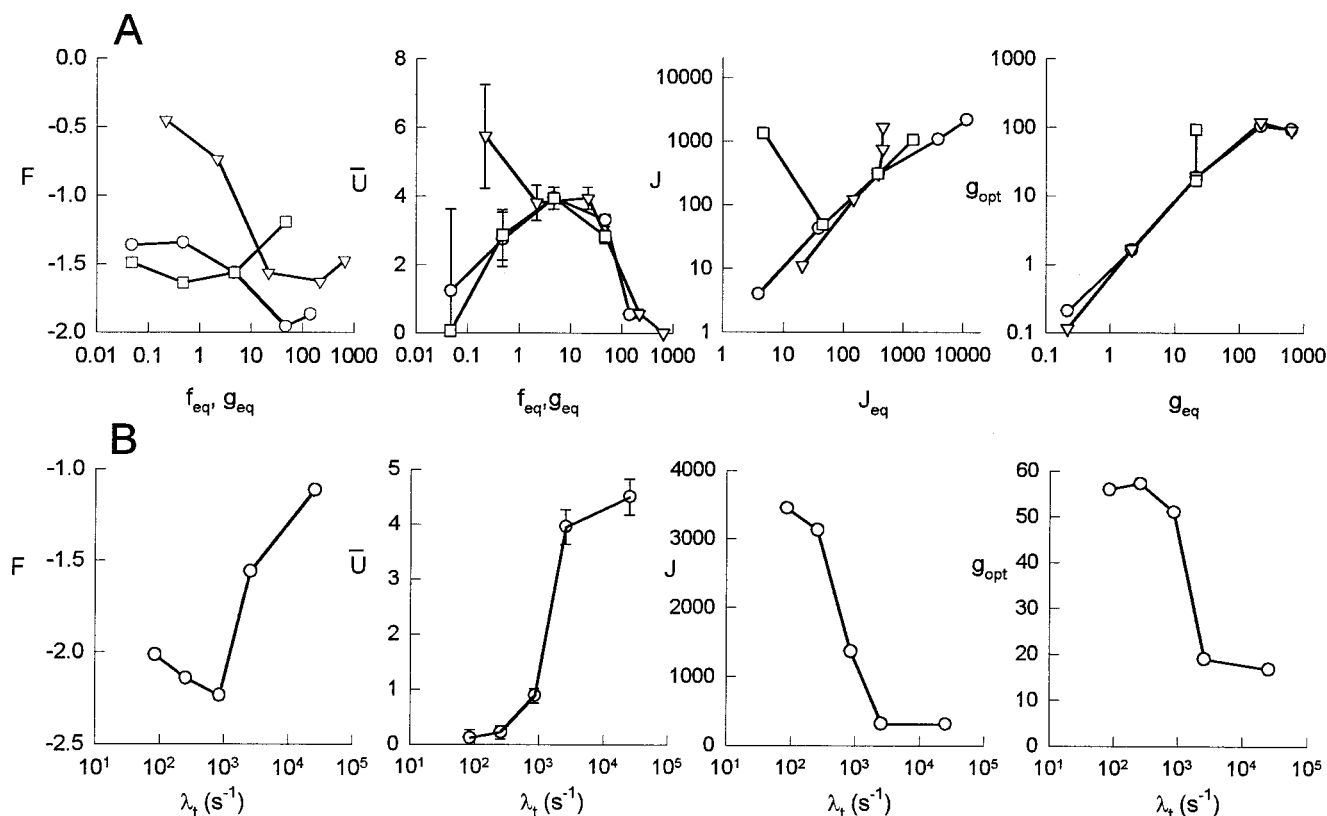
FIGURE 7 Detected output as a function of the characteristic frequencies of events and damping, from 100-s simulations described in the main text. The figure of merit **F**, raw working stroke $\bar{U}$, number of events J, and optimized detachment rate g_{opt} versus predicted attachment rates f_{eq}, g_{eq} (A) and the corner frequency λ_t (B). In A, $\lambda_t = 2500$ s^{-1} and rate constants were varied either at a fixed ratio $f_{eq}/g_{eq} = 0.217$ (○, plotted against f_{eq}), at fixed $f_{eq} = 4.57$ s^{-1} (▽, plotted against g_{eq}) or at fixed $g_{eq} = 21.1$ s^{-1} (□, against f_{eq}). The expected raw working stroke is 4.4 nm. Graphs for J and g_{opt} are plotted against their predicted values $J_{eq} = 100 f_{eq}g_{eq}/(f_{eq} + g_{eq})$ and g_{eq}; there is one failure, on a record containing only six events. In B, both rate constants were fixed as above and the corner frequency was increased beyond the Nyquist critical frequency $\omega_c = 628$ s^{-1}.

ment rate), the performance of the detector can be characterized over a range of simulations. When the simulation model is relevant, the results may be used to guide analysis of experimental data obtained under different conditions, such as ATP level. Simulations at different variance ratios are summarized in Table 2, which confirms conditions (12a and b) in the sense that performance is an increasing function of V_1/V_2 and R.

It remains to vary the rate constants and corner frequency, which limit the sampling rate in the variance record (Eq. 11). For this purpose, simulations were carried out with $h = 5.5$ nm, $V_1/V_2 = 5$, and $\kappa_t = 0.1$ pN/nm, sampled at 10 KHz and varianced with $W = 100$, giving $\Lambda = 5$ ms. The resulting Nyquist frequency $\omega_c = \pi/\Lambda = 628$ s^{-1} is an absolute upper limit for both rate constants in Eq. 11, which implies ATP levels below 0.3 mM for $g = 2 \times 10^6$[ATP] s^{-1}, and ω_c is a lower limit for the corner frequency λ_t.

Fig. 7 shows the chosen indicators as functions of the predicted rate constants f_{eq}, g_{eq} (varied in fixed proportion or singly) and λ_t (varied through the damping constant β). In the same way that log(likelihood)/point cannot be used to compare results from different variance records, the figure

of merit can be used only to indicate the optimum value of W for a given displacement record and not to compare results from different displacement records. This is evident in Fig. 7 A at the highest rates, where **F** (□) > **F** (○) but $\bar{U}$ (□) < $\bar{U}$ (○), and in Fig. 7 B as the corner frequency falls below 1000 s^{-1}, where **F** increases but $\bar{U}$ falls toward zero. The other quantities, particularly the raw working stroke, are useful performance measures across records with different kinetic parameters but the same true working stroke. Under "quasi-rigor" conditions $g_{eq} \to 0$, f_{eq} fixed (Fig. 7 A, plot ▽), the expected level and detachment rate are detected but with low accuracy from the small number of events. Reliable detection fails when the binding frequency is low ($f_{eq} \to 0$, g_{eq} fixed, plot □), or when at least one rate constant rises near or above the Nyquist frequency (628 s^{-1}) of the variance record (Eq. 11). If λ_t falls below 628 s^{-1}, there is an abrupt increase in false detections and an abrupt drop in $\bar{U}$ toward zero (Fig. 7 B), confirming the left-hand inequality in Eq. (11).

Event detection has now been explored as a function of the parameters $c, h, \kappa_t, k, f, g, \Delta, \beta$ of the simulation model. Each displacement-time record can be scaled with respect to

chosen units of length and time, say $\sigma_1 = \sqrt{k_B T / \kappa_t}$ and Δ, leaving the six dimensionless ratios c/σ_1, h/σ_1, k/κ_t, $f\Delta$, $g\Delta$, and $\lambda_t \Delta$ as independent parameters. Quantized bound levels arising from a definite actin site spacing c (either 5.5 or 2.75 nm) are not resolved with the customary 5-nm histogram bins, so the value of c has not been varied. Simulations with the same seed integer but different working strokes h give near-identical running variance records and event detections. The remaining ratios determine the variance ratio and frequency ratios in Eq. 11, whose effects have been described.

Monte Carlo simulations can be made with strain-dependent detachment from bound states, which may accompany ADP release or ATP binding. This effect does enhance the lifetimes of the more positive bound levels, but only weakly because of the averaging effect of free Brownian motions (Smith, 1998b). Displacement events detected from a running-variance record do not easily reveal such correlations (Molloy et al., 1995; Mehta et al., 1997); the correlation would be more marked with stiff traps (Finer et al., 1994), which compromise the variance ratio.

TWO-CHANNEL DATA

The analysis of displacement data from the double-trap experiment (Fig. 1) is often complicated by several possibilities, notably: 1) extendable actin-bead linkages, which are significantly more compliant than the bound myosin; 2) drift in the positions of the optical traps and/or microscope stage; and 3) interactions with more than one myosin. Systematic corrections for the effects of compliant links and trap drift can be made after event detection, provided the x-displacements of both beads are recorded; these claims are backed by Monte Carlo simulations. The third problem is not amenable to correction by analysis. Other problems can arise, but lie beyond the scope of this paper.

Effects of compliant links

The extensibility of actin-bead linkages can be much reduced by pre-tensioning the dumbbell in the traps. The force-extension characteristic of the links is highly nonlinear as the filament peels off the beads (Dupuis et al., 1997). We find that the differential link stiffness can be increased to 1–3 pN/nm (Fig. 8); much lower values are reported in the literature (Mehta et al., 1997; Veigel et al., 1998). The slope of the force-extension curve at the operating point gives the stiffness of the links in series, but compliance corrections require the stiffness of each link over its working range. By using variance information from both bead displacements, effective in situ elastic constants for each link and myosin can be obtained. In practice, the link stiffnesses may be very different. The actin filament is effectively inextensible (Kojima et al., 1994).
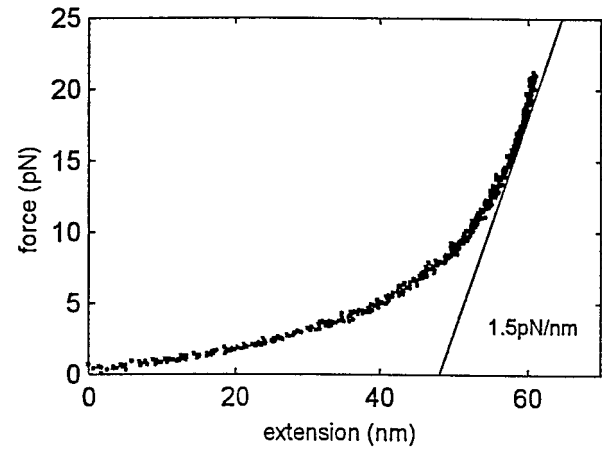

FIGURE 8 The force-extension characteristic of a bead-actin-bead "dumbbell" as measured with 0.085 pN/nm traps. The straight line indicates an achieveable stiffness. The observed nonlinear extensibility is due mainly to the actin-bead links, as the stiffness of our unregulated 6 μm actin filament should be over 7 pN/nm (Kojima et al., 1994).

The method adopted is suggested by the work of Mehta et al. (1997), and requires three variance channels, namely autovariance records for each bead and the covariance record, constructed from running averages $\overline{u_L(t)u_R(t)} - \overline{u_L(t)} \; \overline{u_R(t)}$ from left- and right-bead displacements u_L, u_R. Events should be detected in one autovariance record as before, using the channel with the higher variance ratio (often obvious by inspection). Segmenting the variance records gives the mean variance levels V_{1L}, V_{1R}, V_{1LR} in free periods and V_{2L}, V_{2R}, V_{2LR} in bound periods. The corresponding ratios can then be used to estimate the elastic constants κ_L, κ_R for left and right links and k for myosin as fractions of the stiffness $\kappa_t/2$ of each trap, using the formulae derived in Appendix D:

$$\frac{V_{1L}}{V_{2L}} = \frac{\kappa_L + \tilde{\kappa}_R}{k + \kappa_L + \tilde{\kappa}_R} \frac{k + \tilde{\kappa}}{\tilde{\kappa}} \qquad (19a)$$

$$\frac{V_{1R}}{V_{2R}} = \frac{\tilde{\kappa}_L + \kappa_R}{k + \tilde{\kappa}_L + \kappa_R} \frac{k + \tilde{\kappa}}{\tilde{\kappa}} \qquad (19b)$$

$$\frac{V_{1LR}}{V_{2LR}} = \frac{k + \tilde{\kappa}}{\tilde{\kappa}}. \qquad (19c)$$

Here $\tilde{\kappa}_L = \kappa_t \kappa_L / (\kappa_t + 2\kappa_L)$ is stiffness for the left-hand link and trap, similarly for $\tilde{\kappa}_R$, and $\tilde{\kappa} = \tilde{\kappa}_L + \tilde{\kappa}_R$. Brownian equilibrium is assumed in each attachment period.

Numerical solutions for the stiffness ratios can be obtained by Newton's method (Press et al., 1992) after eliminating Eq. 19c. These ratios also allow the working stroke h to be estimated from mean bound levels, using the formulae

$$\bar{U}_L = \frac{2\tilde{\kappa}_L}{\kappa_t} \frac{kh}{k + \tilde{\kappa}}, \quad \bar{U}_R = \frac{2\tilde{\kappa}_R}{\kappa_t} \frac{kh}{k + \tilde{\kappa}} \qquad (20)$$
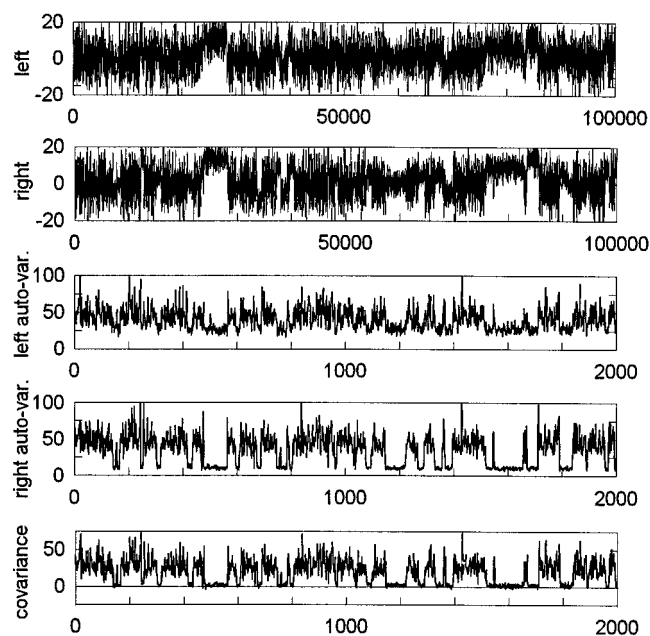
FIGURE 9   Simulated two-channel displacement data (one-tenth of a 100-s record at 10 KHz per channel) and associated running variances for the double-bead system (Fig. 2) with elastically compliant actin-bead links, using $k/\kappa_t = 10$, $\kappa_L/\kappa_t = 1$, $\kappa_R/\kappa_t = 5$, and conventional centered windowing with $W = 100$. The right auto-variance channel was used to detect events. The values of detected parameters (with predicted values in brackets) are $h_L = 5.6 \pm 0.4$, $h_R = 5.5 \pm 0.4$ (5.50) nm, $\bar{U}_L = 3.6 \pm 0.25$ (3.4), and $\bar{U}_R = 4.9 \pm 0.25$ (4.6) nm, $V_{1L} = 46.5$ and $V_{1R} = 46.6$ (49.0), $V_{1LR} = 27.9$ (30.8), $V_{2L} = 27.9$ (28.3), $V_{2R} = 10.8$ (10.3), and $V_{2LR} = 2.21$ (2.25) nm².

obtained by setting $x = 0$ in Eqs. D2b, as the rate of myosin binding peaks at $x = 0$.

To test this procedure, two-channel Monte Carlo simulations were carried out for the elastically coupled bead-filament-bead system, assuming that the filament is lightly damped and in instantaneous mechanical equilibrium with both beads. Fig. 9 shows simulated displacement records and derived variance records for case I ($k = 1.0$ pN/nm, $k/\kappa_t = 10$) but finite and unequal link stiffnesses $\kappa_L = \kappa_t$, $\kappa_R = 5\kappa_t$. The weaker link causes the covariance to drop to very low values in bound periods, as observed experimentally by Mehta et al. (1997), making fluctuations about the mean to positive and negative values. For this reason, the estimated mean bound covariance level is subject to error, which can be reduced by a careful choice of running-variance algorithm. These features confirm that the covariance channel has the biggest free-to-bound variance ratio (Eqs. 19) and the free covariance level is smaller than free autovariance levels (Eqs. D7). Simulations suggest that Eqs. 19 can be generally used to estimate link stiffnesses and compliance corrections.

Correcting for slow drifts

Slow changes in the positions of the traps and the microscope stage with its fixed myosin-coated bead may occur

relative to the quadrant detectors. In the x-direction through the trap centers, the two kinds of drift have different effects on displacement records. Changes in trap position shift the resting position of the filament-beads system, which is defined to be the baseline from which displacements are measured; these movements can be monitored in the absence of a trapped filament and in our experiments are normally <2 nm for 100 s of data acquisition. Drift of the microscope stage on which the fixed bead is mounted may be larger (typically of the order of 10 nm over 100 s once the apparatus is thermally equilibrated). Moving the stage along the x-axis will move all bound levels correspondingly. If all binding sites are orientationally accessible, for example by free rotation of the actin-beads dumbbell during one record, the frequencies at which bound levels appear should be the same Gaussian function of the level (Eq. C1) regardless of the position of the stage; this is the basis of the statistical method for calculating the working stroke (Appendix C).

Running-variance records should not be substantially altered by slow positional drifts of either kind, provided drift is negligible over the windowing time Λ. Hence the positional stability of the traps can be checked in situ by fitting a smooth baseline function $b(t)$ to mean displacement levels in all detected free periods. For this purpose, a chi-squared function that recognizes finite-sampling deviations in each free level is

$$\chi^2 = \sum_{j=1}^{J_2} \frac{(U(j) - b(t(j + 1/2)))^2}{S(j)^2/n(j)} \, \delta_{i(j),1} \qquad (21)$$

where $t(j)$ is the time of the jth event, $i(j)$, $U(j)$, $S(j)$ the state, mean level, and standard deviation in the following period, and $n(j) = t(j + 1) - t(j)$, $t(j + 1/2) \equiv (t(j + 1) + t(j))/2$. It is convenient to use a linear combination of Tschebychev polynomials to order $M < 20$, where M is determined when by the goodness of fit, regressing to a straight-line fit ($M = 1$) if necessary. The fitting routine was tested on Monte Carlo data to which an undulating base line with several turning points and peak-to-peak excursions of up to 20 nm was added. The fitted function $b(t)$ normally deviated from the imposed baseline by 1–2 nm, except near rapidly varying large excursions. To measure displacements from the perceived resting position of the traps, each detected level is corrected by subtracting the value of $b(t)$ at the middle of the attachment period.

The apparent rate of myosin binding is modulated by drift in the vertical positions of the traps relative to the stage. In the presence of a time-varying attachment rate, the variance-HM method detects events without difficulty, as it cannot discriminate between local variations in f and the Poisson distribution of free lifetimes produced by a fixed value of f, 95% of which range over a factor of $6/fW\Delta > 3$ (Eq. 11).

Experimental data

Two-channel displacement data for rabbit myosin-S1 at 5 μM ATP were obtained from a double-beam optical trap system developed in this laboratory. The apparatus is essentially as described by Tskhovrebova et al. (1997), but with a piezoelectric stage and a second quadrant detector. The objective was mounted directly under the stage on a piezoelectric focusing device. The sensitivity of bead movement (volts/nm) was periodically recalibrated by applying a square wave to the piezoelectric transducers carrying the quadrant detectors. The stiffness of the traps was measured either from the amplitude of Brownian motion or by applying a square wave to the acousto-optic modulators and measuring the rate of relaxation of bead position. Displacements of both beads were collected at 10 KHz and written to files of 10^6 points/channel. Data were collected while the $x - y$ position of the microscope stage was electronically stabilized. The apparatus and experimental techniques will also be described elsewhere. All data were analyzed with the variance-HM method.

With fixed trap positions and a stabilized stage, we have found that the distribution of bound levels is obviously non-Gaussian and highly dependent on the x-positions of the traps. Fig. 10 shows two such distributions, with means of 10.0 and 2.3 nm, from single 100-s records with 220 and 286 detected bindings, respectively, using a combined trap stiffness of 0.16 pN/nm and interrelated by a 12-nm displacement of both traps. The reproducibility of these non-Gaussian profiles was demonstrated by analyzing multiple records at a fixed trap position. Studies with weaker traps (0.04 pN/nm) revealed bimodal and occasionally trimodal distributions of bound levels (Simmons et al., 2001). A comprehensive study of this effect shows that the tethered myosin is selecting a subset of monomeric actin sites on a non-rotating actin filament (Steffen et al., submitted for publication). Whatever the cause, it is clear that the conditions for a valid application of Molloy's method for measuring the working stroke are not met (Appendix C); a new experimental protocol is required.

To obtain a Gaussian distribution of bound levels, displacement data were collected when both traps were moved at a constant x-velocity of 0.7–0.8 nm/s by acousto-optic modulators, covering approximately two 36-nm repeats of the actin double-helix over each 100-s recording period. In this way, target zones on each strand of the filament are distributed uniformly in time through the continuous range of x-positions that would be presented by all sites on a freely rotating filament between fixed traps, as envisaged by Molloy et al. (1995). Bead displacements recorded with moving traps as above were analyzed for events using the channel with the higher variance ratio. Baseline fitting was used to reconstruct the applied ramp waveforms in each channel. These waveforms were subtracted from detected displace-
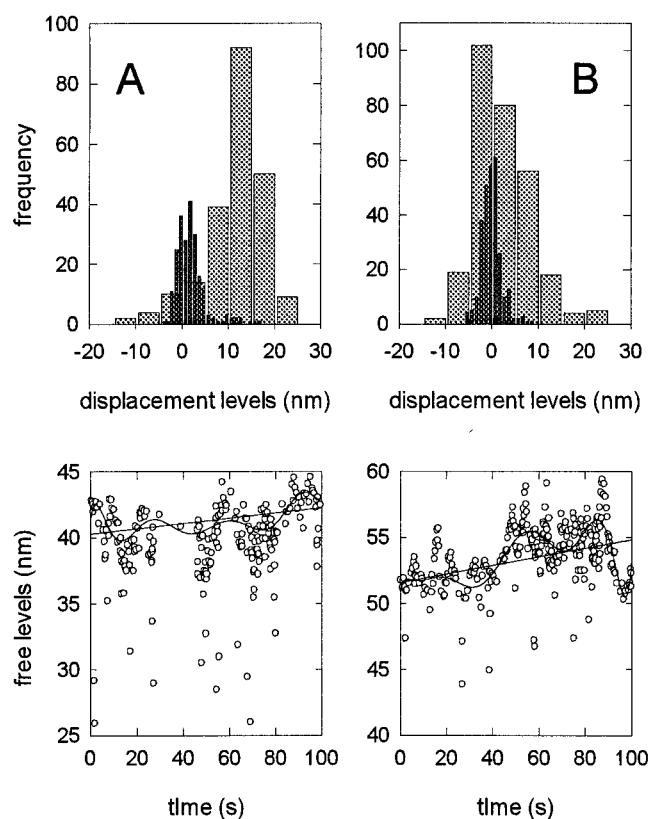

FIGURE 10   Distributions of bound displacement levels (*top figures*) and baseline determinations (*bottom figures*) obtained from the double trap system for myosin S1 at 5 μM ATP, with fixed traps having a combined stiffness of 0.16 pN/nm. Reaction conditions were 25 mM KCl, 25 mM Hepes, 4 mM $MgCl_2$, pH 7.6 and 24°C. A phosphocreatine-creatine kinase ATP regenerating system was used when myosin-S1 had been purified by spinning down degraded heads in the presence of F-actin and ATP. Biotinylated actin was attached to streptavidin beads as described by Ishijima et al. (1998). Plots *A* and *B* were obtained before and after displacing both trap centers by ~20 nm in the x-direction, using acousto-optic modulators. For presentation, baselines determined by the free levels shown were fitted either as a straight line or a 10th-order Tschebychev polynomial. The distribution of bound levels changes slightly with the method of baseline fit, but the mean level is invariant.

ment levels to measure the latter from the resting position of the trap at the time.

Fig. 11 shows a portion of one 100-s record obtained in this way after subtracting the fitted baseline, and the associated running variances, using 0.08 pN/nm traps. Inspection of the whole record (not shown) suggests that the frequency of attachment events is modulated by trap motion, with recurring dense and sparse periods.

Output information for this record is shown in Fig. 12. The distribution of bound levels (Fig. 12 *C*) is essentially the Gaussian distribution of free displacements of the dumbbell, displaced by an uncorrected working stroke of 6.2 $\pm$ 0.4 nm. Using Eqs. 19, both link stiffnesses were estimated at 4 pN/nm, which accounts for the very high variance ratio (~25) achieved in either channel. With such strong links,
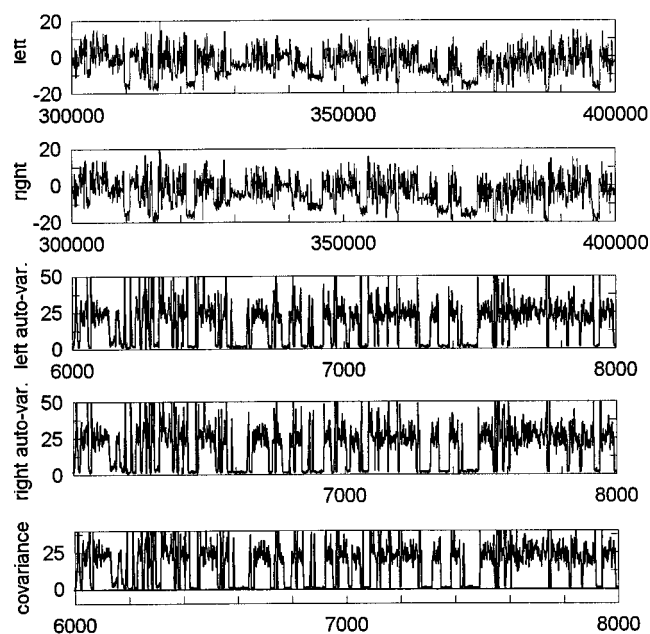
FIGURE 11 Experimental data from the same dumbbell, reaction medium, and trap strengths as in Fig. 10, while the traps were moved at constant velocity in the x-direction. Bead displacements are measured by fixed quadrant detectors and trap movements obtained after event detection by fitting the free periods as described in the text; the resulting straight line fits in each channel have been subtracted to display bead displacements relative to their trap positions. The three running variances shown underneath were constructed at $W = 100$ by conventional centered windowing and clipped at three times their average level.

the similarly large myosin stiffness estimated from covariance levels is probably not reliable; values of 1–2 pN/nm are typical for data obtained with weaker traps. A small compliance correction (5%) is required to obtain the working stroke of the tethered myosin-S1 (6.5 ± 0.4 nm). This error arises from the standard deviations of free and bound levels, not errors in event detection. We also find the expected Poisson distributions of free and bound lifetimes (Fig. 12 D), as observed by Finer et al. (1995).

Fig. 12, A–D were obtained with a window size of 100 points. When analyzed with different window sizes, the estimated working stroke (and the figure or merit with $\varepsilon = 0.0028$) peak near $W = 50$ (Fig. 12 E). However, plotting the detected mean lifetime $1/g_{opt}$ as a function of $\Lambda = W\Delta/2$, as in Fig. 12 F, suggests that $W = 100$ is a more conservative choice; lifetimes obtained with $W > 100$ approximate a straight line of unit slope, which suggests that there are few false detections and that nearly all attachment periods in the variance record are detected (see under Fig. 5). Extrapolation to the ordinate gives $1/g$, where $g = 12.8$ s^{-1} is the detachment rate without windowing, as expected at 5 μM ATP.

The reproducibility of these results was tested over five records obtained with ramped trap movements, using the same dumbbell and fixed bead and, by inference, the same myosin. The stiffnesses of traps and links obtained from detected variance levels varied by 7%, while routine independent measurements showed that the stiffness of the traps was more stable. Table 3 shows that the working strokes in these records have a SD of 0.46 nm, which is little different from the 0.4–0.5-nm finite-sampling variation within each record. Hence the five records can be assumed to sample the same population. Similar results were obtained from the left-hand channel. Averaging over both channels gives an estimated S1 working stroke of 6.3 ± 0.2 nm from this experiment.

The possibility of interactions with more than one myosin on the fixed bead is minimized by coating beads with S1 at dilutions such that interactions can be found on less than half of the fixed beads on the coverslip. The trapped dumbbell was manipulated to search the surface of each fixed bead for interactions; the bead was discarded if the region of interaction was on one side or more than one region was found. The absence of ramped displacements is not a sufficient test for single-S1 interactions, since the filament might interact sequentially with different myosins. In practical terms, a sufficient condition for single-molecule interactions is the observation of a well-defined 5.5-nm periodicity in bound levels from data with fixed trap positions (Steffen et al., submitted for publication), also observed in records analyzed here.

We suggest that it is important to check the validity of the statistical method of Molloy et al. (1995) by fitting the distribution of bound levels to a single Gaussian at the level of a single data record with at least 200 detected events. Non-Gaussian behavior in single records can combine when pooled to produce an apparently Gaussian distribution. A Gaussian distribution of bound levels was also recovered by pooling the data mentioned above with a multiplicity of fixed trap positions spanning an integral number of 36-nm actin repeats. Three such experiments, sampling 3, 7, and 6 repeats at ~10-nm intervals, gave estimated working strokes of 5.4, 5.8, and 5.8 nm. Taken as a whole, the four experiments returned values between 5.4 and 6.3 nm for the working stroke of our tethered myosins.

USING THE METHOD

The variance-Hidden-Markov technique as implemented in this paper is a very robust and effectively "on-line" tool, extensively used on experimental data from our laboratory, for detecting actin-myosin attachment events in displacement records. Output information includes the mean displacements in free and bound periods, and their standard errors from finite sampling; with a Gaussian distribution of bound levels, this information is sufficient to estimate the myosin working stroke. With a 1 GHz computer, the processing time for a binary data record of 10^6 points/channel is ~5 s.
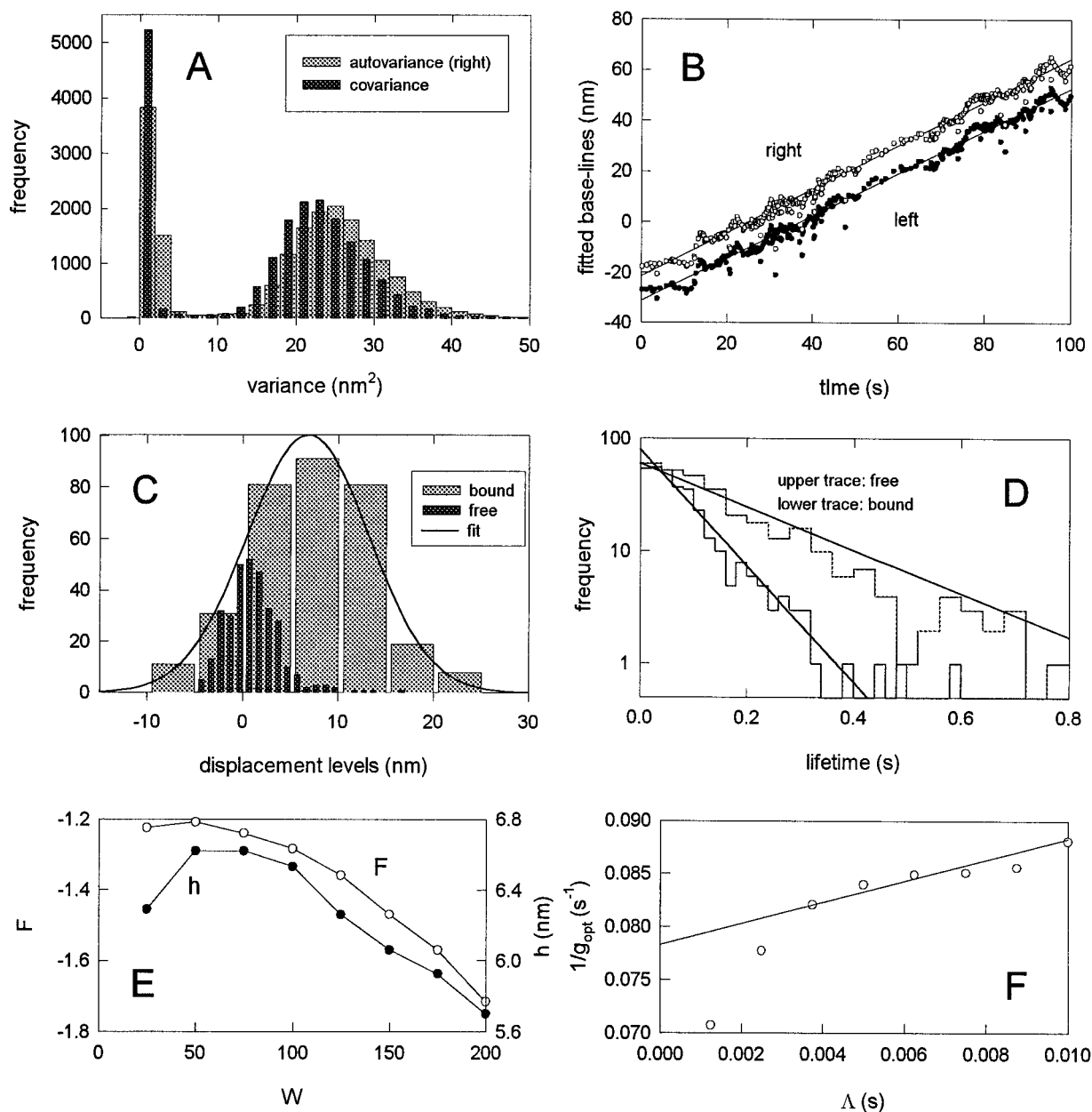
FIGURE 12 Event detection and analysis of the 100-s displacement record with ramped traps, sampled in Fig. 10. (*A*) Distributions of autovariance (right channel) and covariance in the three-channel variance record before detection, showing bimodal behavior from free and bound periods. The fractional half-width of the free mode is ~0.24 (predicted half-width $\sqrt{2/W} = 0.14$ for $W = 100$). (*B*) Trap movements over the record as constructed after event detection by fitting straight lines through displacement levels in free periods ($\bigcirc$ and $\bullet$). (*C*) Distributions of free and bound displacement levels of the right-hand bead measured relative to its trap, and the Gaussian distribution of bound levels that best fits the latter (mean = 6.40 ± 0.24 nm, SD = 6.85 ± 0.37 nm, $\chi^2/N = 0.18$) and a narrower distribution of free levels with some tailing on the bound side. (*D*) Lifetimes of free and bound periods, plotted logarithmically to show the approach beyond the first two bins to Poisson distributions with lifetimes $1/f_{opt}$, $1/g_{opt}$ where $f_{opt} = 4.43$ s^{-1}, $g_{opt} = 11.9$ s^{-1}. (*E, F*) Optimized values of the figure of merit **F**, estimated working stroke h, and mean lifetime $1/g_{opt}$ versus window size W. The mean lifetime is plotted against the sampling time $\Lambda = 5 \times 10^{-5}$ W (s) in the variance record to assess the prediction under Fig. 5; the straight line shown has unit slope and is intended merely to guide the eye.

The variance-HM method is able to detect binding/detachment steps of all sizes, since it is based on a derived running-variance record rather than raw displacement data. However, the limitations of the method should be kept in mind. Only binding and detachment events are detected, not post-binding events such as ADP release, which would not change system stiffness or the bound variance level. Information in the displacement record is lost by the need to

TABLE 3 Variability in detected output from five experimental myosin-S1 records of 10^6 sample points at 5 μM ATP, namely attachment rates *f, g*, the compliance-corrected working strokes *h*, spread of bound levels *S*, and link stiffnesses in each channel (L or R), also the combined stiffness of the traps

J_1, J_2	f (s^{-1})	g (s^{-1})	h_L, h_R (nm)	S_L, S_R (nm)	κ_L, κ_R (pN/nm)	κ_t (pN/nm)
322	4.43	11.90	6.16 ± 0.39	6.08	4.26	0.163
			6.53 ± 0.41	6.38	3.92	
323	4.34	12.34	6.26 ± 0.44	6.55	3.46	0.163
			6.60 ± 0.45	6.68	3.77	
224	2.74	12.20	5.90 ± 0.46	5.89	4.31	0.168
			6.23 ± 0.47	6.03	4.00	
238	2.93	12.56	6.75 ± 0.46	6.24	3.99	0.164
			7.09 ± 0.48	6.43	3.80	
215	2.63	11.55	5.46 ± 0.42	5.29	3.97	0.175
			5.70 ± 0.43	5.50	4.07	
Mean	3.57	12.1	6.13 ± 0.20	6.09	3.98	0.166
			6.46 ± 0.20	6.29	3.90	

All elastic constants are estimated from mean variance levels as described in text and Appendix D. The quoted standard errors arise only from finite sampling. Records were combined using Eqs. 16–18 of the text. The spreads of bound levels is ~1 nm more than as predicted by the kinetic model of Appendix D, namely $\sqrt{k_B T/\kappa_t} = 4.9$ nm, which is consistent with residual drift in the position of the microscope stage.

construct running-variances averaged over windows of 50 points or more. The method works only within a certain range of trap stiffnesses (typically 0.01–0.5 pN/nm). If trap stiffness approaches or exceeds the stiffness of myosin or the actin-bead links, detection becomes difficult as the ratio of free to bound variance levels is reduced toward unity (Eqs. 12); in these circumstances a Hidden-Markov method is required that operates directly on displacement data. As the traps are weakened, the corner frequency of the dumbbell is lowered and the method fails when this frequency falls below the sampling frequency of the variance record (Eq. 11). However, the sampling frequency must obviously be greater than event frequencies, including the ATP-dependent detachment rate. Hence the method works well over a range of ATP concentrations limited from above by the corner frequency of the free dumbbell and from below by the stability of the experimental system over the time required to detect many events.

Within these limitations, our preferred strategy for using the variance-HM method is as follows. Event detection should initially be carried out with a window of 200 points ($W = 200$), and accepted only if the detected variance ratio is above, say, four in at least one channel, which should then be used for detection. The optimal window size W_{opt} should be sought by maximizing the figure of merit **F** provided that the raw working stroke, if any, is also maximized at a similar value of W. The detected detachment rate and number of events should fall steadily as W is increased above the optimum value. Unacceptable behavior as W is decreased below W_{opt} is signaled by a decrease in apparent working

stroke and a sudden increase in detected events, irrespective of the change in **F**. The number of detections is usually reflected in the detachment rate *g*. Independent knowledge of *g* from solution-kinetic data can serve to identify unrealistically high detected values, which indicate false detections. The variance-HM method may also return an unexpectedly low working stroke if the myosin is loosely attached to the coverslip, which is indicated by a low returned value for myosin stiffness. In practice, these checks appear sufficient to validate the method within the above experimental constraints.

Supplementary material describing our implementation of the variance-Hidden Markov method, and executable program files and source code, can be downloaded from www.kcl.ac.uk/depsta/biomedical/randall/mrcmcmu.html. The authors take no responsibility for the accuracy or end use of programs provided in this way.

CONCLUDING DISCUSSION

Experimental constraints on the correct use of the variance-HM method have already been discussed. We now compare this method with other methods for analyzing displacement data, and assess the origins of the different estimates from various laboratories for the myosin working stroke.

Hidden-Markov methods provide a globally optimizing form of event detection by re-estimating the four parameters *f, g*, V_1, V_2 of the underlying two-dimensional stochastic process (Appendix A), and the window width W. They are superior to single-event detectors, whose operation is necessarily threshold-dependent (Pastushenko and Schindler, 1993; Smith, 1998a), because the likelihood will rise when threshold constraints are removed (Appendix B), for example by using the variance-HM method. The nature of the improvement is determined by the quality of the data. With good running-variance data (say $V_1/V_2 > 5$, mean lifetimes $>8\Lambda$ for $W = 100$), nearly all detected events detected by the variance-HM method pass Student's *t*-test at the 5% confidence level, which sets a threshold signal relative to variance "noise" (Eq. 5); in this case the results should be relatively insensitive to changing the confidence level and the Hidden-Markov method should not make a significant improvement. When these criteria are not met, "single-event" detectors may be threshold-sensitive; whether this is so for the corpus of published data with V_1/V_2 in the range 2–5 and $W < 100$ is not clear.

The possibility of a drifting baseline raises questions about how working strokes should be measured before making compliance corrections. If the baseline is a known function of time, sampled working strokes are obtained by subtracting this function from all bound levels. However, a fitted baseline could be quite inaccurate in some parts of the record. As an alternative to global fitting, the set of levels in all free periods can be taken to <u>define</u> a baseline function

(Veigel et al., 1998) after suitable interpolation. This procedure ignores Brownian noise, so it is instructive to compare the two methods. Over 10 experimental records, the two methods produced essentially working strokes differing by 0.2 or 0.45 nm at most in each record. This remarkable agreement is a result of the large number of sampled displacements in each free period.

Design principles for running-variance algorithms follow by considering the intrinsic bandwidth of a running-variance record and how it can best be sampled. The bandwidth (in radians/s) of a variance record constructed continuously in time by sweeping a window of W data points through a displacement record sampled at intervals of Δ is basically of the order of $2\pi/W\Delta$, although the precise high-frequency roll-off of the power spectrum depends on the algorithm. This frequency is also the Nyquist critical frequency ω_c when the variance record is sampled at multiples of $\Lambda = W\Delta/2$; hence little information will be gained by more frequent sampling, even when followed by resampling filters with a final bandwidth of $2\pi/W\Delta$. Second, it might appear advantageous to construct running-variances with a much smaller window $W' \ll W$ (to increase time resolution) followed by edge-preserving filtering down to a bandwidth of ω_c (to reduce fluctuations). No amount of post-filtering can improve the poor quality of variance records constructed in this way, which show high noise and an abnormally low apparent variance ratio compared to our standard method (sampling at intervals of $W\Delta/2$) with the same final bandwidth.

Comparisons with the mean-variance method (Patlak, 1993; Guilford et al., 1997) are instructive. Patlak's method does not seek to identify individual attachment events, but to observe bound displacement levels in a limited class of distributions (running means and variances of point displacements). The method is optimizing for the distributions above, and rates of binding and detachment are estimated by fitting the distributions over a range of window sizes. Thus all information required to construct the transition matrix is generated, but this matrix is not used to make stochastic predictions. The need to fit mean-variance distributions over a large range of window widths W may make it less accurate than the variance-HM method, where fully optimized parameter fitting is available for any value of W. However, it cannot be concluded that the prominent bound levels at ± 11 nm observed by Guilford et al. (1997) are an artefact of the mean-variance method; we have also observed a bimodal distribution of bound levels (Simmons et al., 2001) with traps of similar strength (0.03–0.04 pN/nm), which can be explained in helical selection of actin sites on a non-rotating filament. The mean-variance method was also used by Tyska et al. (1999), who do not observe a bimodal distribution.

In most studies, actin-bead linkages have been much weaker than myosin stiffness as currently estimated (1–2 pN/nm), so that compliance corrections to the working stroke will be greater than the 5% correction for our data with 4 pN/nm links. We suggest that in situ stiffness measurements should be used to make compliance corrections, even when they have been minimized by pre-tensioning the dumbbell. The discrepant working strokes reported by different groups might arise from compliance corrections and/or a non-Gaussian distribution of bound levels. Without more experimentation, it is hard to say whether the latter arise from inadequate sampling or the distortions caused by a non-rotating filament. If data from contiguous records is pooled, the effects of a non-rotating filament should generally be mitigated by longitudinal drift in the microscope stage; however, the "ramped-traps" technique introduced in this paper should generally be more accurate. The orientation of the tethered myosin may be influenced by its mode of attachment to the fixed bead and is not controlled in our experiments. The question of the orientation dependence of the working stroke is not resolved; measurements using a synthetic myosin filament system (Tanaka et al., 1998; Ruff et al., 2001) give conflicting answers.

Estimates of errors arising from the variance-HM method do not normally include errors in event detection. This defect can be rectified within a fully Bayesian approach, as described by Rosales et al. (2001). Events can also be detected by neural networks or wavelet analysis; although these methods are very powerful they are less specific because no assumptions need be made about the nature of the stochastic process.

Variance-HM detection is widely applicable because it requires almost no information about the actin-myosin crossbridge cycle. The method is indifferent to the number of accessible binding sites on the filament, the degree of access, or bound states on the same site that generate different amounts of force without changing the stiffness of the crossbridge. Changes in stiffness occurring after myosin binding, or by the binding of a second myosin, would probably not be detected even if the variance-HM method were generalized to include a third variance level. Orientationally disordered bound states might generate little or no extra stiffness and might not be detected. In particular, the method does not require that myosin make a working stroke after binding.

APPENDIX A

A general Hidden-Markov method

Consider a sequence $X \equiv (x_1, \ldots, x_N)$ of N data points, collected at times $t_k = k\Delta$, $k = 1, \ldots, N$. The physical interpretation of the x_k varies according to the application, and need not be specified. We assume that X is stochastically determined by a two-dimensional Markov process operating on the data in conjunction with a set of discrete states $i = 1, \ldots, M$ hidden from the observer. This process is specified by compound transition probabilities $A_{ij}(x_k, x_{k+1})$ where $A_{ij}(x_k, x_{k+1})dx_{k+1}$ is the probability of state j and data-point in the range $(x_{k+1}, x_{k+1} + dx_{k+1})$ at time t_{k+1}, given state i and data-point x_k at time t_k. The initial probability $P_i(x_1)$ for state i and the first data-point x_1 is also required. In the classical version of the

method, the transition matrix is assumed to be independent of the initial data-value x_k (Rabiner, 1989).

By varying the parameters of the transition matrix, the stochastic model is "fitted" to the data by maximizing the probability P that the model generates the observed time series:

$$P = \sum_{i_1,\ldots,i_N} P_{i_1}(x_1) A_{i_1 i_2}(x_1, x_2) \cdots A_{i_{N-1} i_N}(x_{N-1}, x_N). \quad (A1)$$

P is a generalization of the likelihood function of multivariate statistics (Morrison, 1978) to include a hidden Markov process. The right-hand side may be computed iteratively from partial sum-products $\alpha_{k,i}$ by Baum's forward algorithm (Rabiner, 1989 and references therein)

$$\alpha_{k+1,j} = \sum_i \alpha_{k,i} A_{ij}(x_k, x_{k+1}) \quad (A2a)$$

$$\alpha_{1,i} = P_i(x_1), \quad P = \sum_j \alpha_{N,j} \quad (A2b)$$

not displaying their dependence on the data-values. A similar procedure running backward from the end produces conjugate partial quantities $\beta_{k,i}$ where

$$\beta_{k,i} = \sum_j A_{ij}(x_k, x_{k+1}) \beta_{k+1,j} \quad (A3a)$$

$$\beta_{N,i} = 1, \quad P = \sum_i P_i(x_1) \beta_{1,i}. \quad (A3b)$$

Now $\alpha_{k,i}\beta_{k,i}/P$ is the probability of state i at time t_k, and $\alpha_{k,i}A_{ij}(x_k, x_{k+1})\beta_{k+1,j}/P$ the joint probability of states i, j at times t_k and t_{k+1}. With the observed data, the expected number of transitions from state i to state j is just

$$\xi_{ij} = \frac{1}{P} \sum_{k=1}^{N-1} \alpha_{k,i} A_{ij}(x_k, x_{k+1}) \beta_{k+1,j}. \quad (A4)$$

Rate constants for transitions between the discrete states of the Markov process are embodied by the "states-only" transition matrix

$$a_{ij} = \iint A_{ij}(x, x') \varphi_i(x) dx dx' \quad (A5)$$

where $\varphi_i(x)$ is the normalized probability distribution of data-value x in state i. The transition rates implied by this quantity can be re-estimated from the ξ_{ij} by a procedure due to Baum (1972), giving new "states-only" transition probabilities

$$\tilde{a}_{ij} = \frac{\xi_{ij}}{\sum_j \xi_{ij}} \quad (A6)$$

which update the matrix a_{ij}, and hence the state-dependence of the full transition matrix $A_{ij}(x, x')$. As both quantities are conditional probabilities, they must be normalized to unity over their final states:

$$\sum_j \int A_{ij}(x, x') dx' = 1, \quad \sum_j a_{ij} = 1 \quad (A7a, b)$$

where (A7b) follows from (A7a) and Eq. A5. The importance of Baum's re-estimation method is that it immediately determines the values of the a_{ij} that maximize P. Baum's re-estimation theorem is proved in Appendix B; proofs in terms of a Liapunov function have been given by Baum (1972) and Levinson et al. (1983).

In general, P should also be maximized with respect to other parameters of the transition matrix. It is convenient to carry out Baum's kinetic re-estimation procedure at every step of a full optimization routine with respect to the remaining parameters, because kinetic re-estimation is fast: in the absence of round-off error the optimum rate constants are achieved at the first iteration when the remaining parameters are fixed (Appendix B). In practice, a few Baum iterations are usually required to achieve convergence.

Values of $\alpha_{k,i}$ and $\beta_{k,i}$ calculated from the forward and backward algorithms often fall below computational limits and must be rescaled when necessary. Let the forward recursion (A2a) be written in vector form as $\alpha_{k+1} = \alpha_k \mathbf{A}_k$, where the matrix $\mathbf{A}_k$ has elements $A_{ij}(x_k, x_{k+1})$. This recursion can be replaced by a scaled version

$$\alpha^*_{k+1} = 10^{-e_{k+1}} \alpha^*_k \mathbf{A}_k, \quad (A8a)$$

where $e_{k+1} = \text{int}\{\log_{10}(\max \alpha^*_k \mathbf{A}_k)\}$ with $\alpha^*_1 = \alpha_1$ and $e_1 = 0$. These exponents also generate scaled-back vectors

$$\beta^*_k = 10^{-e_{k+1}} \mathbf{A}_k \beta^*_{k+1} \quad (A8b)$$

so $\alpha^*_k = 10^{-r_k}\alpha_k$, $\beta^*_k = 10^{-s_k}\beta_k$, where $r_k = e_1 + \cdots + e_k$ and $s_k = e_{k+1} + \cdots + e_N$. Thus $r_k + s_k = E \equiv e_1 + \cdots + e_n$ for all time points k. Hence the scaled likelihoods P^* calculated from forward or backward iterations with scaled vectors should be the same and equal to $10^{-E} P$. The quantities ξ_{ij} are invariant under scaling.

If the transition matrix favors a single reaction pathway, the likelihood function could be approximated by the greatest single term in each sum over states. This is the basis of Viterbi's algorithm, which requires only MN rather than M^2N operations (Rabiner, 1989; Patlak, 1997). Whatever method is used, it is useful to record the most probable state at each time point k after optimization, obtained as the value of i for which the re-estimated quantity

$$\tilde{p}_{k,i} = \frac{\alpha_{k,i}\beta_{k,i}}{P} \quad (A9)$$

is a maximum.

APPENDIX B

Baum's re-estimation theorem

Re-estimation, as defined by Eqs. A4 and A6, increases the likelihood function $P(X)$ when the transition matrix describes a coupled two-dimensional Markov process, rather than the classical case of a single hidden Markov process and a continuous random variable without memory. Following Levinson et al. (1983), we show that the re-estimated $P(X)$ is extremal (a maximum or minimum).

Since the likelihood function (Eq. A1) is a sum-over-states of a product of transition matrices and an initial distribution, variations in these quantities give

$$\delta P(X) = \sum_I \left\{ \sum_{k=1}^{N-1} \frac{\delta A_{i_k i_{k+1}}(x_k, x_{k+1})}{A_{i_k i_{k+1}}(x_k, x_{k+1})} + \frac{\delta P_{i_1}(x_1)}{P_{i_1}(x_1)} \right\} P_I(X) \quad (B1)$$

where $I = (i_1, \ldots, i_N)$ and $P_I(X)$ is the summand of (A1). Equivalently,

$$\delta P(X) = \sum_I \left\{ \sum_{k=1}^{N-1} \frac{\delta a_{i_k i_{k+1}}}{a_{i_k i_{k+1}}} + \frac{\delta p_{i_1}}{p_{i_1}} \right\} P_I(X) \quad (B2)$$

where $p_i = \int P_i(x) dx$ is the initial probability of state i. The form of the data distributions is assumed to be unaffected by these variations. They are

constrained by normalization conditions

$$\sum_j a_{ij} = 1, \quad \sum_i p_i = 1. \qquad (B3)$$

On using the identity $P_1(X) = \alpha_{k,i} A_{ij}(x_k, x_{k+1}) \beta_{k+1,j}$ for any k, where i and j are the states at steps k, $k + 1$ respectively, and using undetermined multiplers, we have the extremal condition

$$\sum_{ij} \left\{ \sum_{k=1}^{N-1} \alpha_{k,i} A_{ij}(x_k, x_{k+1}) \beta_{k+1,j} - \lambda_i a_{ij} \right\} \frac{\delta a_{ij}}{a_{ij}}$$

$$+ \sum_i (\alpha_{1,i} \beta_{1,i} - \lambda p_i) \frac{\delta p_i}{p_i} = 0. \quad (B4)$$

Hence

$$\lambda_i a_{ij} = \sum_{k=1}^{N-1} \alpha_{k,i} A_{ij}(x_k, x_{k+1}) \beta_{k+1,j},$$

$$\lambda p_i = \alpha_{1,i} \beta_{1,i} \qquad (B5)$$

so

$$\lambda_i = \sum_j \sum_{k=1}^{N-1} \alpha_{k,i} A_{ij}(x_k, x_{k+1}) \beta_{k+1,j},$$

$$\lambda = \sum_i \alpha_{1,i} \beta_{1,i} \qquad (B6)$$

on using Eqs. B3. Hence the solutions of Eqs. B5 are

$$\tilde{a}_{ij} = \frac{\displaystyle\sum_{k=1}^{N-1} \alpha_{k,i} A_{ij}(x_k, x_{k+1}) \beta_{k+1,j}}{\displaystyle\sum_j \sum_{k=1}^{N-1} \alpha_{k,i} A_{ij}(x_k, x_{k+1}) \beta_{k+1,j}} \qquad (B7)$$

and

$$\tilde{p}_i = \frac{\alpha_{1,i} \beta_{1,i}}{\displaystyle\sum_i \alpha_{1,i} \beta_{1,i}}. \qquad (B8)$$

Equation B7 determines a new "states-only" transition matrix as in Eq. A6, and Eq. B8 an updated initial state probability; this is the content of Baum's re-estimation theorem. In general, this probability will not be equal to that calculated ab initio from the updated transition matrix. For ergodic stochastic problems, the effects of updating the initial state should not propagate in time, so updating the initial state probability merely generates an end-effect of $O(1/N)$ for the total likelihood.

The above argument shows that the likelihood function is extremal after one re-estimation. That the extremum is in fact a maximum is proved by Baum (1972), and Levinson et al. (1983). Conditions for the existence of a global maximum were investigated by Baum et al. (1970).

APPENDIX C

The statistical method for determining the working stroke

Molloy et al. (1995) have identified the most frequently accessed bound displacement level of the actin-bead system (Fig. 1) as the uncorrected myosin working stroke. They observe a Gaussian distribution of binding frequencies $J(U)$ by level, namely

$$J(U) \propto \exp\left(-\frac{(U - \bar{U})^2}{2\sigma^2}\right) \qquad (C1)$$

where the standard deviation $\sigma \approx \sigma_1 \equiv \sqrt{k_B T / \kappa_t}$ reflects Brownian fluctuations of the filament before binding. Then the working stroke is determined by the mean displacement level $\bar{U}$ and Eq. 2. The justification for this procedure is statistical, as summarized under Fig. 1.

Equations 2 and C1 can be derived for an adaptation of A. F. Huxley's crossbridge model (Huxley, 1957) to multiple binding sites on the actin filament. This demonstration requires that the rate constant $f(x)$ for myosin binding to a site at distance x on the stationary filament (Fig. 1) is localized within a narrow range about zero strain, here defined as $x = 0$. An implicit force-generating transition after binding (Geeves and Holmes, 1999) is implied if the final force is $k(x + h)$, which displaces the actin filament by an amount $k(x + h)/(k + \kappa_t)$ (Smith, 1998b).

This model generates a discrete set of bound displacement levels if myosin can bind within a cluster of monomeric sites on each half-pitch of the actin double-helix. Let the positions of sites in the nearest cluster be $x_l \equiv x + lc$, where l is an integer and c the monomer spacing. When the filament-bead system is weakly trapped, its longitudinal Brownian displacements should be sufficient to present all these sites to the tethered myosin, particularly if the filament undergoes torsional oscillations (Svensson and Thomas, 1986) or uniform rotation about an axis through the traps (Molloy et al., 1995). Hence the Brownian-averaged binding rate to site l (Smith, 1998b) is

$$\bar{f}_\ell = \int f(x_\ell - u) \frac{1}{\sqrt{2\pi\sigma_l^2}} \exp(-u^2/2\sigma_1^2) du. \qquad (C2)$$

Under fairly general conditions, this integral reduces to an expression of the form

$$\bar{f}_\ell \approx A \exp(-x_\ell^2/2\sigma_s^2). \qquad (C3)$$

If $f(x)$ is supported over a range $\ll \sigma_1$ about $x = 0$, then C3 applies with $\sigma_s \approx \sigma_1$ and $F = (2\pi\sigma_1^2)^{-1/2} \int f(x)dx$. If $f(x) = F \exp(-x^2/2\sigma_M^2)$ where $\sigma_M = \sqrt{k_B T/k}$ is the thermal reach of myosin with stiffness k, the integral in C2 evaluates exactly to C3 with $A = (1 + k/\kappa_t)^{-1/2}F$ and $\sigma_s^2 = \sigma_1^2 + \sigma_M^2$. In the second case, the binding range σ_s is determined by the stiffness of myosin acting in series with both traps, reducing when $k \gg \kappa_t$ to the first case as required.

The number of binding events J_ℓ per second to site l is determined from the kinetic equations of a multi-site Huxley model with Brownian-averaged rate constants $\bar{f}_\ell$, $\bar{g}_\ell$. Under steady-state conditions $J_\ell = \bar{f}_\ell(1 - p)$, where p is the overall probability of attachment. For a single molecule, this formula predicts the average rate of binding (and detachment) over many attachment events, over which steady-state conditions should prevail. Thus

$$J_\ell = C \exp(-x_\ell^2/2\sigma_s^2) \qquad (C4)$$

where $C = A(1 - p)$. With weak traps ($k \gg \kappa_t$), the site-dependence of J_ℓ reflects the Boltzmann frequency distribution of Brownian displacements of the free filament. As the ATP level is decreased, all binding rates decrease uniformly, $p \to 1$ and $C \to 0$. However, kinetic models valid for cycling conditions may not apply in this limit.

As the bound displacement level for site l is

$$U(x_\ell) = \frac{k(x_\ell + h)}{k + \kappa_t} \qquad (C5)$$

the binding rate J_ℓ can be viewed as a function of displacement level rather than the site index l. In this sense, the multi-site kinetic model leads to the desired result (C1) with a standard deviation

$$\sigma = \frac{k\sigma_s}{k + \kappa_t},\qquad\text{(C6)}$$

also of order σ_1 for weak traps. This model also predicts quantized levels that vary linearly with x (the position of the nearest actin site), but these features are obscured if the binding rate is assumed to be a continuous function of U.

The rate constants of the variance-HM model describe binding and detachment without regard to actin site. Equivalent rate constants f_{eq} and g_{eq} may therefore be defined from the multi-site model as

$$f_{eq} = \sum_\ell \bar{f}_\ell, \quad g_{eq} = \frac{1}{p}\sum_\ell p_\ell \bar{g}_\ell \qquad\text{(C7)}$$

where

$$p_\ell = \frac{\bar{K}_\ell}{1 + \sum_\ell \bar{K}_\ell} \quad \left(\bar{K}_\ell \equiv \frac{\bar{f}_\ell}{\bar{g}_\ell}\right) \qquad\text{(C8)}$$

is the occupation probability of the bound state on site l and $p = \Sigma_\ell\, p_\ell$. These definitions are consistent with detailed balancing, as $f_{eq}(1 - p) = g_{eq}p$.

APPENDIX D

Displacement and variance levels with compliant actin-bead links

The mechanical system shown in Fig. 1 allows an exact calculation of the mean displacements and variances of the beads under conditions of thermal equilibrium. Bead displacements are relative to their resting positions without bound myosin, and variances are required for both attachment states. Fluctuating displacements are generated primarily by Brownian forces on the beads, but their (Boltzmann) distribution in equilibrium is governed only by the absolute temperature T and the potential energy function

$$V(u_L, u_R, u) = \frac{\kappa_t}{4}(u_L^2 + u_R^2) + \frac{\kappa_L}{2}(u - u_L)^2 + \frac{\kappa_R}{2}$$

$$\times (u - u_R)^2 + \frac{k}{2}(x + h - u)^2 \qquad\text{(D1)}$$

for bound myosin. Here u_L, u_R, u are the left-directed displacements of the left-hand and right-hand beads and the filament; other symbols are defined in the figure caption. Results for free myosin follow by setting $k = 0$.

Conditions for mechanical equilibrium ($\partial V/\partial u_L = 0$, etc.) give the resting positions of the filament and beads, namely

$$U(x) = \frac{k(x + h)}{k + \tilde{\kappa}_L + \tilde{\kappa}_R},\qquad\text{(D2a)}$$

$$U_L(x) = \frac{\kappa_L U(x)}{\frac{1}{2}\kappa_t + \kappa_L}, \quad U_R(x) = \frac{\kappa_R U(x)}{\frac{1}{2}\kappa_t + \kappa_R} \qquad\text{(D2b)}$$

where

$$\tilde{\kappa}_L = \frac{\frac{1}{2}\kappa_t\kappa_L}{\frac{1}{2}\kappa_t + \kappa_L}, \quad \tilde{\kappa}_R = \frac{\frac{1}{2}\kappa_t\kappa_R}{\frac{1}{2}\kappa_t + \kappa_R} \qquad\text{(D3)}$$

are elastic constants for each trap and link in series. These resting positions are also the mean displacements in the presence of Brownian forces when the system has come to thermal equilibrium, with a Boltzmann distribution of displacements. Using a vector notation for the column vector (u_L, u_R, u), the mean displacement levels are

$$\underline{U} = \frac{\displaystyle\int \underline{u}\,\exp(-V(\underline{u})/k_BT)d\underline{u}}{\displaystyle\int \exp(-V(\underline{u})/k_BT)d\underline{u}} \qquad\text{(D4)}$$

which reproduces Eqs. D2 as the potential is harmonic. If the force-extension relation for traps or links were nonlinear, the mean equilibrium positions of the beads would be different from their resting positions.

In the same way, variances are evaluated as components of the thermal average of the matrix $\underline{\delta u}\,\underline{\delta u}^T$, where $\underline{\delta u} = \underline{u} - \underline{U}$ is a column vector and $\underline{\delta u}^T$ the corresponding row vector. The quantity

$$\mathbf{V} = \frac{\displaystyle\int \underline{\delta u}\,\underline{\delta u}^T\exp(-V(\underline{u})/k_BT)d\underline{u}}{\displaystyle\int \exp(-V(\underline{u})/k_BT)d\underline{u}} \qquad\text{(D5)}$$

is a symmetric 3×3 matrix of variances. The potential energy is a symmetric quadratic function $V(\underline{u}) = \underline{\delta u}^T \mathbf{A}\underline{\delta u}/2$ of deviations from the mean, where

$$\mathbf{A} = \begin{pmatrix} \dfrac{\kappa_t}{2} + \kappa_L & 0 & -\kappa_L \\[2mm] 0 & \dfrac{\kappa_t}{2} + \kappa_R & -\kappa_R \\[2mm] -\kappa_L & -\kappa_R & \kappa_L + \kappa_R + k \end{pmatrix} \qquad\text{(D6)}$$

The denominator in Eqs. D4 and D5 is equal to $(2\pi k_BT)^{3/2}|A|^{-1/2}$ (Doi and Edwards, 1988), so $\mathbf{V} = k_BT\, d(\ln|\mathbf{A}|)/d\mathbf{A} \equiv RT\, \mathbf{A}^{-1}$. The left and right autovariances V_{LL}, V_{RR}, and the covariance V_{LR} can be read from the upper-diagonal 2×2 partition, giving

$$V_{LL} = k_BT\,\frac{k + \kappa_L + \tilde{\kappa}_R}{\left(\dfrac{\kappa_t}{2} + \kappa_L\right)(k + \tilde{\kappa}_L + \tilde{\kappa}_R)} \qquad\text{(D7a)}$$

$$V_{RR} = k_BT\,\frac{k + \kappa_R + \tilde{\kappa}_L}{\left(\dfrac{\kappa_t}{2} + \kappa_R\right)(k + \tilde{\kappa}_R + \tilde{\kappa}_L)} \qquad\text{(D7b)}$$

$$V_{LR} = k_BT\,\frac{\kappa_L\kappa_R}{\left(\dfrac{\kappa_t}{2} + \kappa_L\right)\left(\dfrac{\kappa_t}{2} + \kappa_R\right)(k + \tilde{\kappa}_L + \tilde{\kappa}_R)} \qquad\text{(D7c)}$$

for the bound levels, and for the free levels with $k = 0$. The quantities $\tilde{\kappa}_L$, $\tilde{\kappa}_R$ are defined in the main text. As expected, the free autovariances for each channel are the same: this is not immediately apparent in the formulae.

These results yield Eqs. 11 of the main text for the three ratios of free-bound variance levels. Setting $x = 0$ in Eqs. D7 gives Eqs. 12 for bound displacement levels. The argument for setting $x = 0$ (Appendix C) is kinetic. Note also two special cases in which the results simplify: 1) when κ_L, $\kappa_R \gg k$, κ_t the autovariances for each channel satisfy $V_1/V_2 = 1 + k/\kappa_t$ and the mean bound displacement level $\bar{U}$ is given by Eq. 2 (the rigid-dumbbell model). 2) If the links are equally stiff ($\kappa_L = \kappa_R = \kappa/2$), $\bar{U}$ is given by a modification of Eq. 2 with myosin stiffness k replaced by the stiffness $\tilde{k} = k\kappa/(k + \kappa)$ of myosin in series with both links (Smith, 1998b). However, autovariance levels in each attachment state do not correspond to the same modification of Eq. 1; Eqs. D7 give autovariance levels

$$\frac{V_1}{k_B T} = \frac{1}{\kappa_t} + \frac{1}{\kappa + \kappa_t}, \quad \frac{V_2}{k_B T} = \frac{1}{\tilde{k} + \kappa_t} + \frac{1}{\kappa + \kappa_t} \quad \text{(D8a)}$$

for each bead, while the covariance levels for free and bound myosin are

$$\frac{V_{1LR}}{k_B T} = \frac{1}{\kappa_t} - \frac{1}{\kappa + \kappa_t}, \quad \frac{V_{2LR}}{k_B T} = \frac{1}{\tilde{k} + \kappa_t} - \frac{1}{\kappa + \kappa_t}. \quad \text{(D8b)}$$

The autovariances exceed those in Eq. 1 with $k \to \tilde{k}$. With myosin free, the covariance is smaller than $k_B T/\kappa_t$. With myosin bound, the covariance is much smaller than the autovariance, particularly if the links are weak, but still positive as $\tilde{k} < \kappa$.

Where comparisons can be made, our results agree with those of Veigel et al. (1998).

REFERENCES

Ball, F. G., and M. S. P. Sansom. 1989. Ion-channel gating mechanisms: model identification and parameter estimation from single channel recordings. *Proc. R. Soc. Lond. B.* 236:385–416.

Baum, L. E. 1972. An inequality and associated maximization technique in statistical estimation for probabilistic functions of a Markov process. *Inequalities.* 3:1–8.

Baum, L. E., T. Petrie, G. Soules, and N. Weiss. 1970. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Stat.* 41:164–171.

Block, S. M., and K. Svoboda. 1995. Analysis of high-resolution recordings of motor movement. *Biophys. J.* 68:230S–241S.

Chung, S. H., V. Krishnamurthy, and J. B. Moore. 1991. Adaptive processing techniques based on Hidden Markov Models for characterising very small channel currents buried in noise and deterministic inferences. *Phil. Trans. R. Soc. Lond. B.* 334:357–384.

Chung, S. H., J. B. Moore, L. Xia, L. S. Premkumar, and P. W. Gage. 1990. Characterization of single channel currents using digital signal processing techniques based on Hidden Markov Models. *Phil. Trans. R. Soc. Lond. B.* 329:265–285.

Colquhoun, D. 1998. Single-Channel Recording. E. Sakmann and E. Neher, editors. Plenum, New York.

Doi, M., and S. F. Edwards. 1988. Theory of Polymer Dynamics. Oxford University Press.

Dupuis, D. E., W. H. Guilford, J. Wu, and D. M. Warshaw. 1997. Actin filament mechanics in the laser trap. *J. Muscle Res. Cell Motil.* 18:17–30.

Feller, W. 1957. An Introduction to Probability Theory and its Applications. Vol. I, 2nd Ed. John Wiley and Sons, New York and London.

Finer, J. T., A. D. Mehta, and J. A. Spudich. 1995. Characterization of single actin-myosin interactions. *Biophys. J.* 68:291S–297S.

Finer, J. T., R. M. Simmons, and J. A. Spudich. 1994. Single myosin molecule mechanics: piconewton forces and nanometer steps. *Nature.* 368:113–119.

Fredkin, D. R., and J. A. Rice. 1992. Maximum likelihood estimation and identification directly from single-channel recordings. *Proc. R. Soc. Lond. B.* 249:125–132.

Geeves, M. A., and K. C. Holmes. 1999. Structural mechanism of muscle contraction. *Annu. Rev. Biochem.* 68:687–728.

Guilford, W. H., D. E. Dupuis, G. Kennedy, J. Wu, J. B. Patlak, and D. M. Warshaw. 1997. Smooth and skeletal muscle myosins produce similar unitary forces and displacements in the laser trap. *Biophys. J.* 72:1006–1021.

Horn, R., and K. Lange. 1983. Estimating kinetic constants from single channel data. *Biophys. J.* 43:207–223.

Huxley, A. F. 1957. Muscle structure and theories of contraction. *Prog. Biophys. Biophys. Chem.* 7:255–318.

Huxley, A. F., and R. M. Simmons. 1971. Proposed mechanism of force generation in striated muscle. *Nature.* 233:533–538.

Huxley, A. F., and S. Tideswell. 1996. Filament compliance and tension transients in muscle. *J. Muscle Res. Cell Motil.* 17:507–551.

Ishijima, A., H. Kojima, T. Funatsu, H. Higuchi, H. Tanaka, and T. Yanagida. 1998. Simultaneous observation of individual ATPase and mechanical events by a single myosin molecule during interaction with actin. *Cell.* 92:161–171.

Kishino, A., and T. Yanagida. 1988. Force measurements by micromanipulation of a single actin filament by glass needles. *Nature.* 334:74–76.

Kojima, H., A. Ishijima, and T. Yanagida. 1994. Direct measurement of stiffness of single actin filaments with and without tropomyosin by in vitro nanomanipulation. *Biophys. J.* 91:12962–12966.

Kron, S. J., and J. A. Spudich. 1986. Fluorescent actin filaments move on myosin fixed to a glass surface. *Proc. Natl. Acad. Sci. U.S.A.* 83:6272–6276.

Levinson, S. E., L. R. Rabiner, and M. M. Sondhi. 1983. An introduction to the application of the theory of probabilistic functions of a Markov process to automatic speech recognition. *Bell System Tech. J.* 62:1035–1074.

Lorenz, M., K. J. V. Poole, D. Popp, G. Rosenbaum, and K. C. Holmes. 1995. An atomic model of the unregulated thin filament obtained by x-ray fiber diffraction on oriented actin-tropomyosin gels. *J. Mol. Biol.* 246:108–119.

Lymn, R. W., and E. W. Taylor. 1971. Mechanism of adenosine triphosphate hydrolysis by actomyosin. *Biochemistry.* 10:4617–4624.

Mehta, A. D., J. T. Finer, and J. A. Spudich. 1997. Detection of single molecule interactions using correlated thermal diffusion. *Proc. Natl. Acad. Sci. U.S.A.* 94:7927–7931.

Mehta, A. D., and J. A. Spudich. 1998. Single myosin molecule mechanics. *Adv. Struct. Biol.* 5:229–270.

Miyata, H., H. Hakozaki, H. Yoshikawa, N. Suzuki, K. Kinosita, T. Nishizaka, and S. Ishiwata. 1994. Stepwise motion of an actin filament over a small number of heavy meromyosin molecules is revealed in an in vitro motility assay. *J. Biochem. (Tokyo).* 115:644–647.

Molloy, J. E., J. E. Burns, J. Kendrick-Jones, R. T. Tregear, and D. C. S. White. 1995. Movement and force produced by a single myosin head. *Nature.* 378:209–212.

Morrison, D. F. 1978. Multivariate Statistical Analysis, 2nd Ed. McGraw-Hill Int., New York.

Pastushenko, V. Ph., and H. Schindler. 1993. Statistical filtering of the single ion channel records. *Ann. Pharamacol.* 43:7–13.

Patlak, J. B. 1993. Measuring kinetics of complex single ion-channel data using mean-variance histograms. *Biophys. J.* 65:29–42.

Patlak, J. B. 1997. Combined use of mean-variance analysis and the Viterbi algorithm to idealize single channel and single myosin records. *Biophys. J.* 72:327a. (Abstr.).

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. R. Flannery. 1992. Numerical Recipes in Fortran, 2nd Ed. Cambridge, U.K.

Rabiner, L. R. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE.* 77:257–285.

Rosales, R., J. A. Stark, W. J. Fitzgerald, and S. B. Hladky. 2001. Bayesian restoration of ion channel records using Hidden Markov models. *Biophys. J.* 80:1088–1103.

Ruff, C., B. Brenner, and E. Meyerhöfer. 2001. The step size of myosin is orientationally independent. *Biophys. J.* 80:78a. (Abstr.).

Simmons, R. M., D. A. Smith, and J. Sleep. 2001. Target zones on the actin filament and the myosin working stroke from optical trapping. *Biophys. J.* 80:80a. (Abstr.).

Smith, D. A. 1998a. A quantitative method for the detection of edges in noisy time series. *Phil. Trans. R. Soc. B.* 353:1969–1981.

Smith, D. A. 1998b. Direct tests of muscle crossbridge theories: predictions of a Brownian dumbbell model for position-dependent crossbridge lifetimes and step sizes with an optically trapped actin filament. *Biophys. J.* 75:2996–3007.

Svensson, E. C., and D. D. Thomas. 1986. ATP induces microsecond rotational motions of myosin heads crosslinked to actin. *Biophys. J.* 50:999–1002.

Svoboda, K., and S. M. Block. 1994. Biological applications of optical forces. *Annu. Rev. Biophys. Biomol. Struct.* 23:247–285.

Tanaka, H., A. Ishijima, M. Honda, K. Saito, and T. Yanagida. 1998. Orientation dependence of displacements by a single one-headed myosin relative to the actin filament. *Biophys. J.* 75:1886–1894.

Tskhovrebova, L., J. Trinick, J. A. Sleep, and R. M. Simmons. 1997. Elasticity and unfolding of the giant muscle protein titin. *Nature.* 387:308–312.

Tyska, M. J., D. E. Dupuis, W. H. Guilford, J. B. Patlak, G. S. Waller, K. M. Trybus, D. M. Warshaw, and S. Lowey. 1999. Two heads of myosin are better than one for generating force and motion. *Proc. Natl. Acad. Sci. U.S.A.* 96:4402–4407.

Veigel, C., M. L. Bartoo, D. C. S. White, J. C. Sparrow, and J. E. Molloy. 1998. The stiffness of rabbit skeletal actomyosin cross-bridges determined with an optical tweezers transducer. *Biophys. J.* 75:1424–1438.

Veigel, C., L. M. Coluccio, J. D. Jontes, J. C. Sparrow, R. A. Milligan, and J. E. Molloy. 1999. The motor protein myosin-I produces its working stroke in two steps. *Nature.* 398:530–533.

Weatherburn, C. E. 1968. A First Course in Mathematical Statistics, 2nd Ed. Cambridge University Press.